Quantification of Epistemic and Aleatory Variability: Towards Removing the Ergodic Component of Site Response

Adrian Rodriguez-Marek and Mahdi Bahrampouri

Virginia Tech

USA

October 9, 2018

Outline

- Background on non-ergodic PSHA
- Quantification of aleatory variability
- Epistemic uncertainty in the site term
  - Empirical quantification
  - Numerical estimates
- Virginia Tech Sigma2 Research:
  - Project plan
  - Progress to date

PSHA

- A proper PSHA must account for all sources of uncertainty:

UNCERTAINTIES

ALEATORY	**EPISTEMIC**
Random variability [within parametric space]	Lack of knowledge
Accounted for in the *Hazard Integral*	**Accounted for through *Logic-trees***

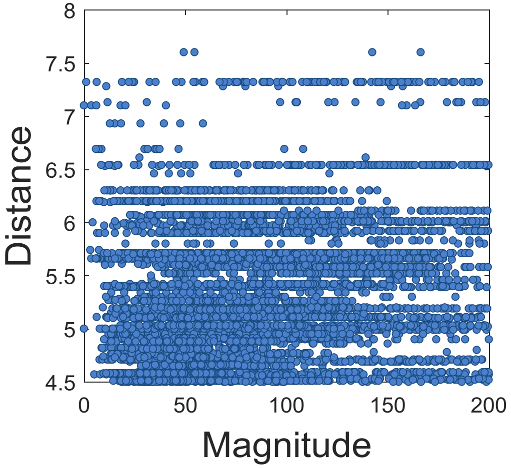
- Proper separation is important
  - Epistemic uncertainty is reflected in *fractiles* of hazard
  - Study of hazard fractiles identifies inputs that can be lead to hazard reduction

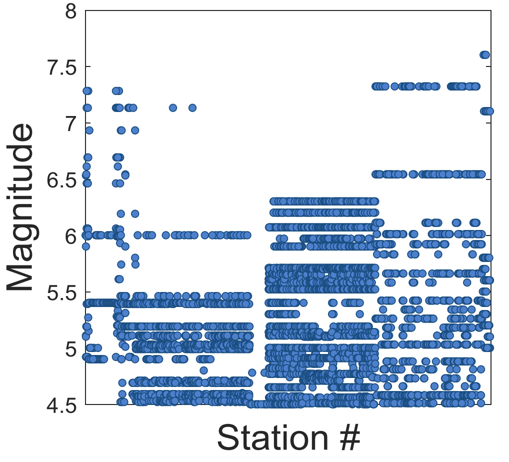
Background on non-ergodic PSHA

$$lnY_{es} = f(M, R, V_{S30}, \dots) + \Delta_{es}$$

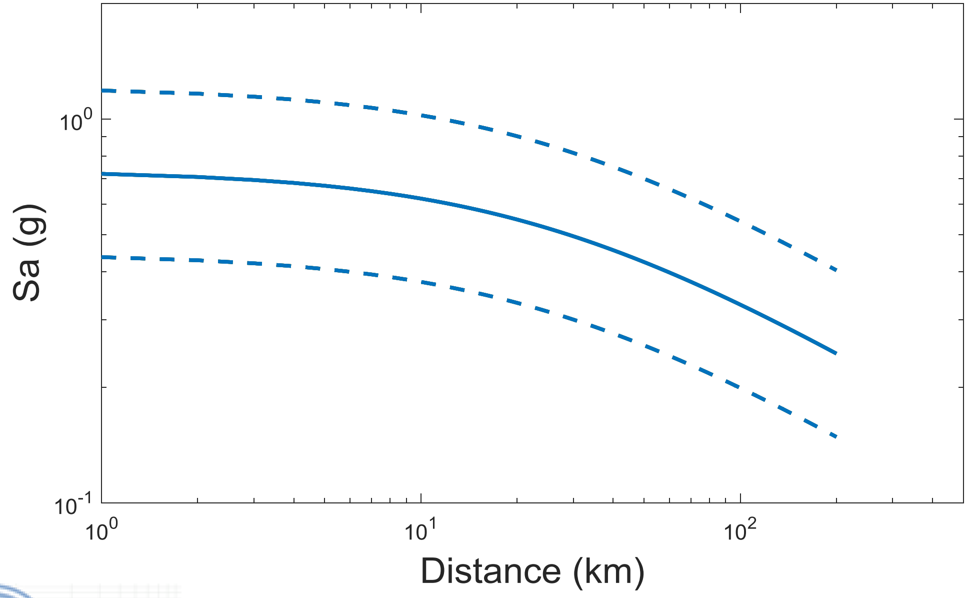
Median

Residual

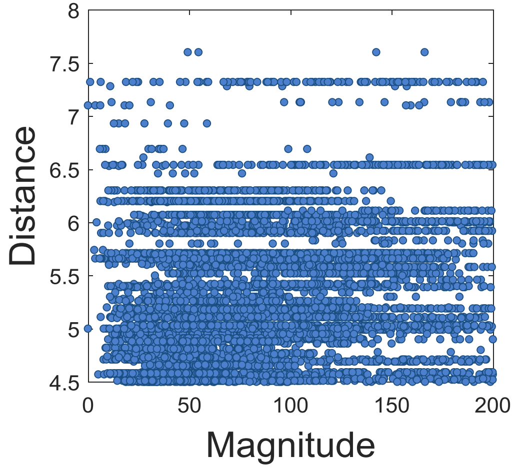


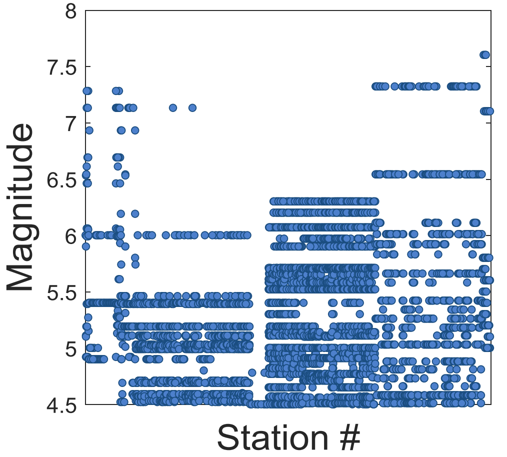


Regression



Background on non-ergodic PSHA





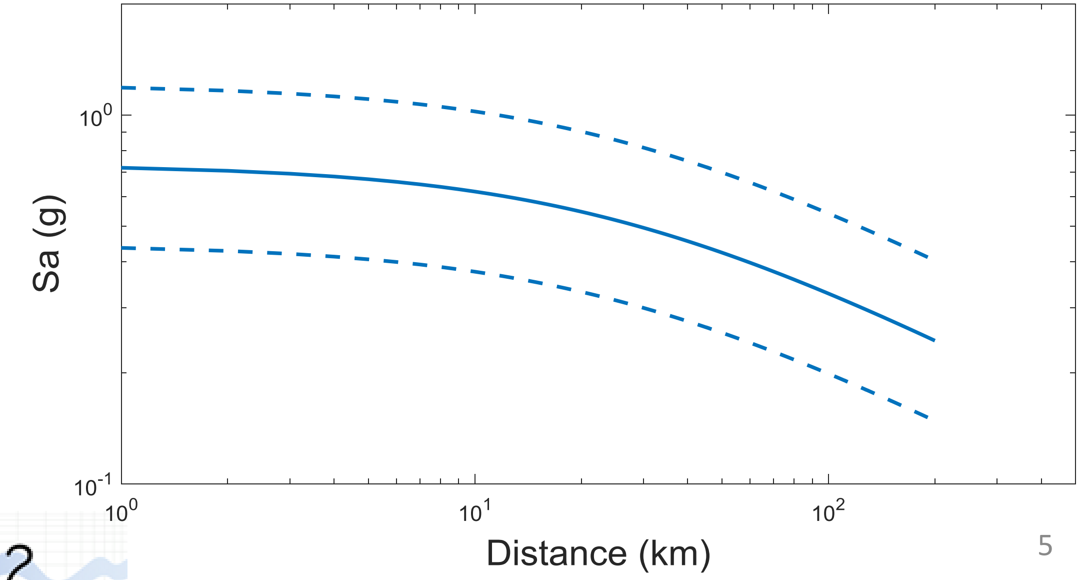
Random effects regression

$$lnY_{es} = f(M, R, V_{S30}, \dots) + \delta B_e + \delta W_{es}$$

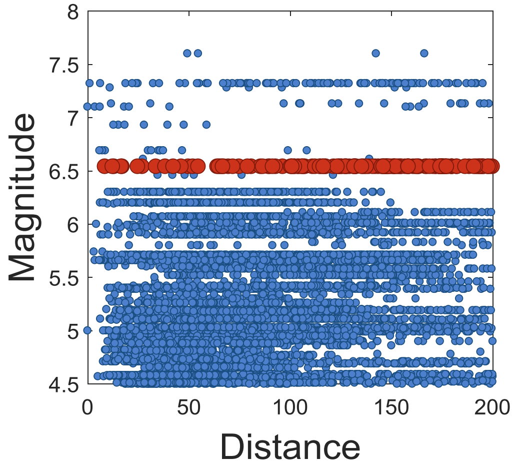
Median

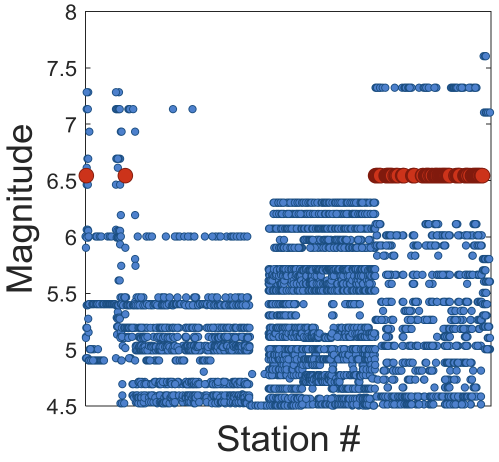
Between event Residual

Within event Residual



Background on non-ergodic PSHA





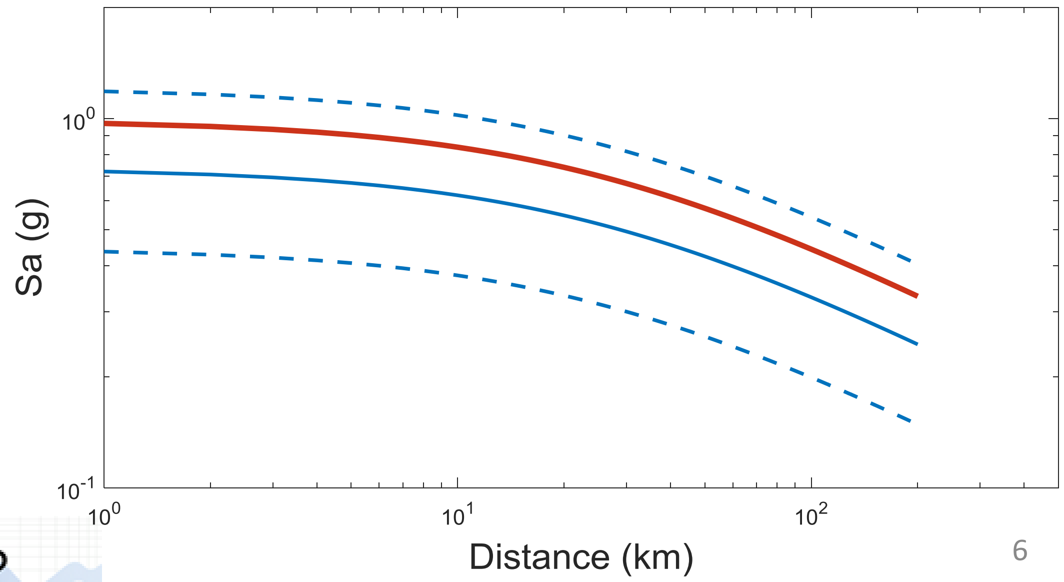
Random effects regression

$$lnY_{es} = f(M, R, V_{S30}, \ldots) + \boldsymbol{\delta B_e} + \delta W_{es}$$

Median

Between event Residual (event term)

Within event Residual



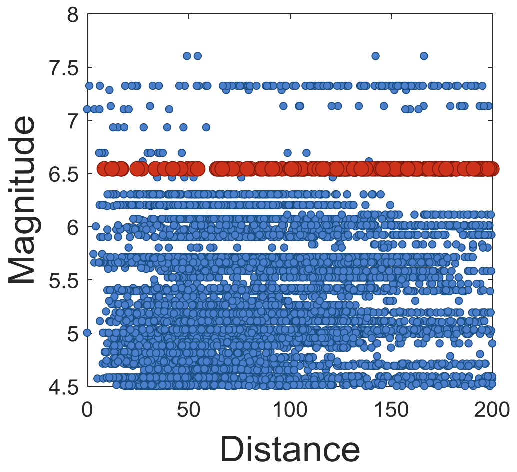
Background on non-ergodic PSHA

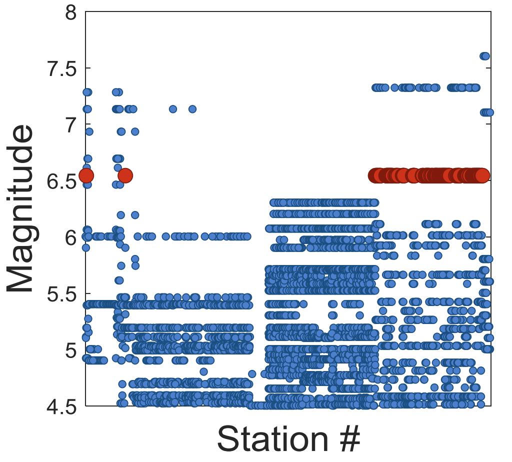
$$lnY_{es} = f(M, R, V_{S30}, \dots) + \boldsymbol{\delta B_e} + \delta W_{es}$$

Median

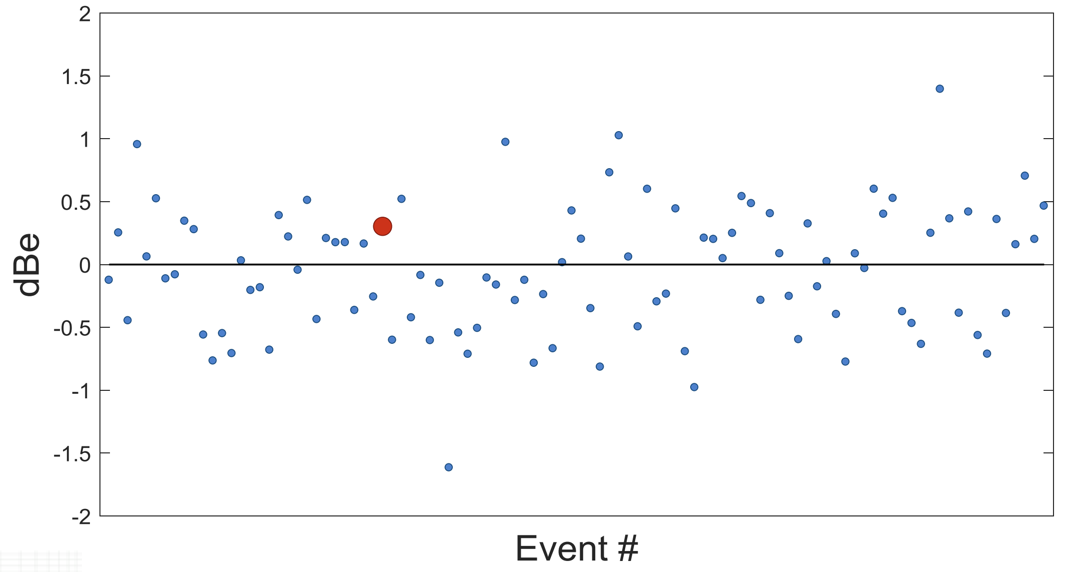
Between event Residual (event term)

Within event Residual





Random effects regression



Background on non-ergodic PSHA

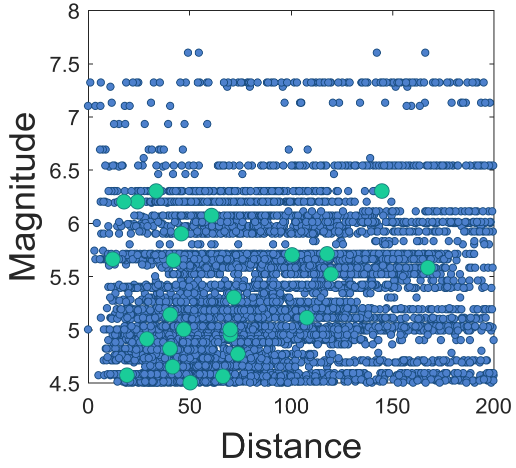
$$lnY_{es} = f(M, R, V_{S30}, \ldots) + \boldsymbol{\delta B_e} + \boldsymbol{\delta S_s} + \delta_o$$

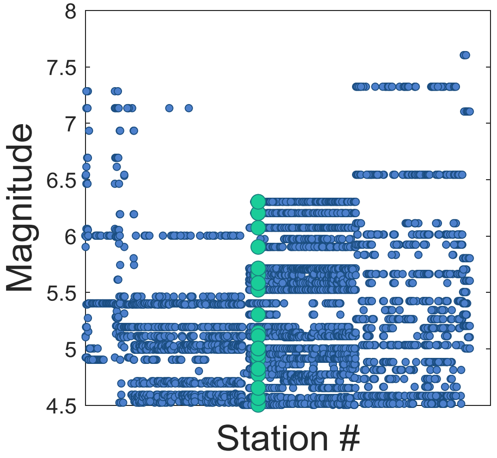
Median

Event Term

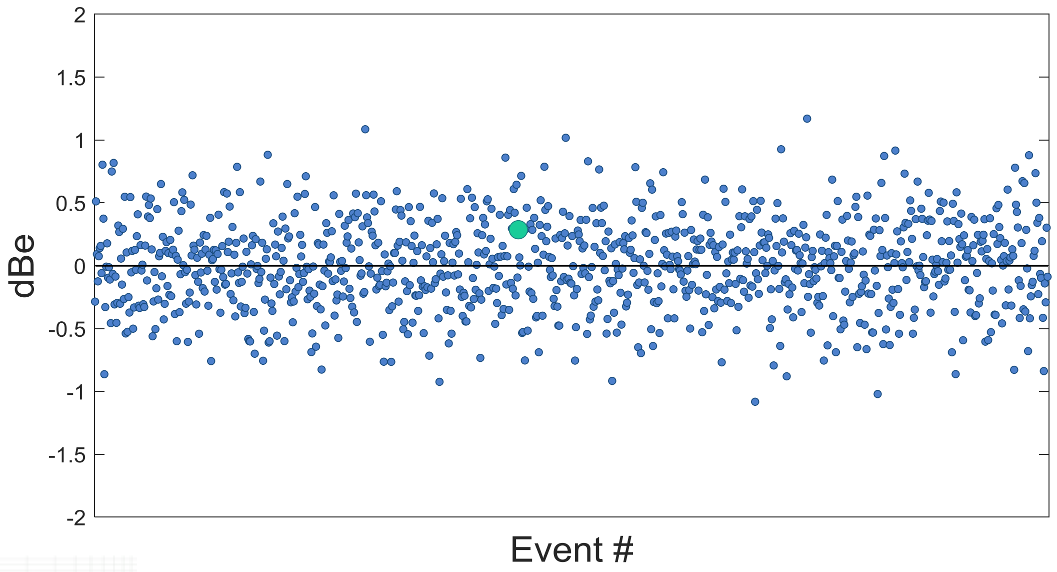
Site Term

Residual





Mixed effects regression



Background on non-ergodic PSHA

- In general:

$$lnY = f(M, R, V_{S30}, \dots) + \delta L_l + \delta P_{sl} + \delta S_s + \delta_{esl}$$

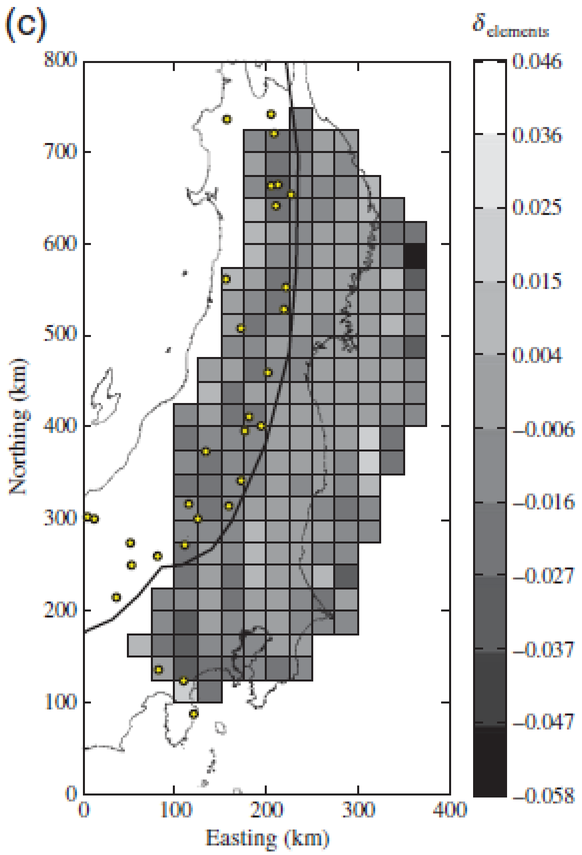
Fixed Effects Random Effects Residual

- Random Effects represent effects that are repeatable
  - δL_l Repeatable source effect (one term for each source location l)
  - δP_{sl} Repeatable path effect (one term for path between source l and site s)
  - δS_s Repeatable site effect (one term for each site)

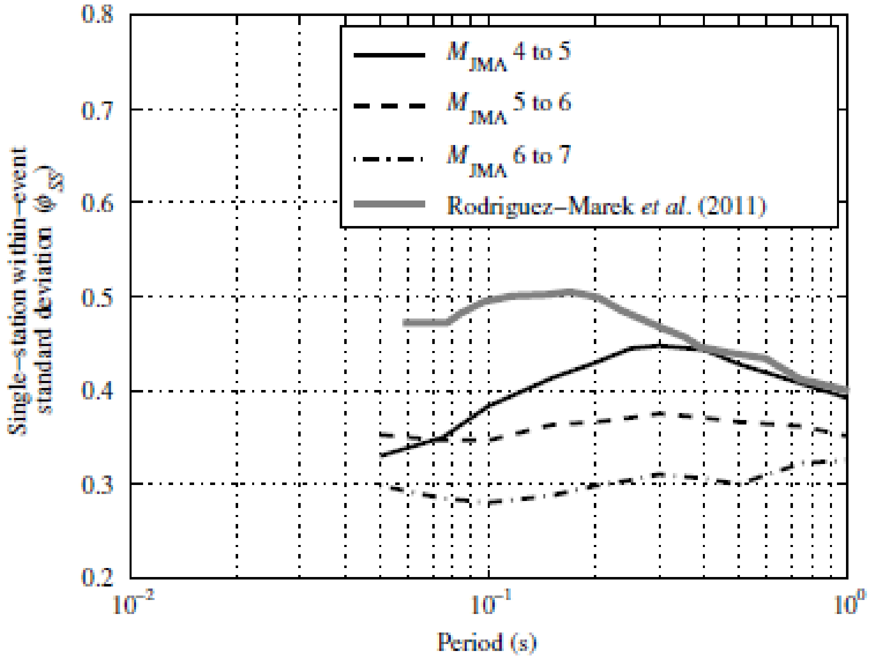
Background on non-ergodic PSHA

- Repeatable effects are, in principle, predictable
- If they are predictable, then uncertainty about these factors should be part of the epistemic uncertainty
- Example: removing path effects in the KiKnet database (Sigma project, Dawood and Rodriguez-Marek, 2013)

$$lnY = \overline{lnY} + \sum_{i=1}^{n} \delta_i P_{ies}$$



Path term for $6 < M \leq 7$ EQs



Reduction in standard deviation

Background on non-ergodic PSHA

- In our Sigma2 research, we focus on a removing the ergodicity related to the site term

$$lnY = f(M, R, V_{S30}, \dots) + \delta S_s + \delta B_e + \delta_0$$

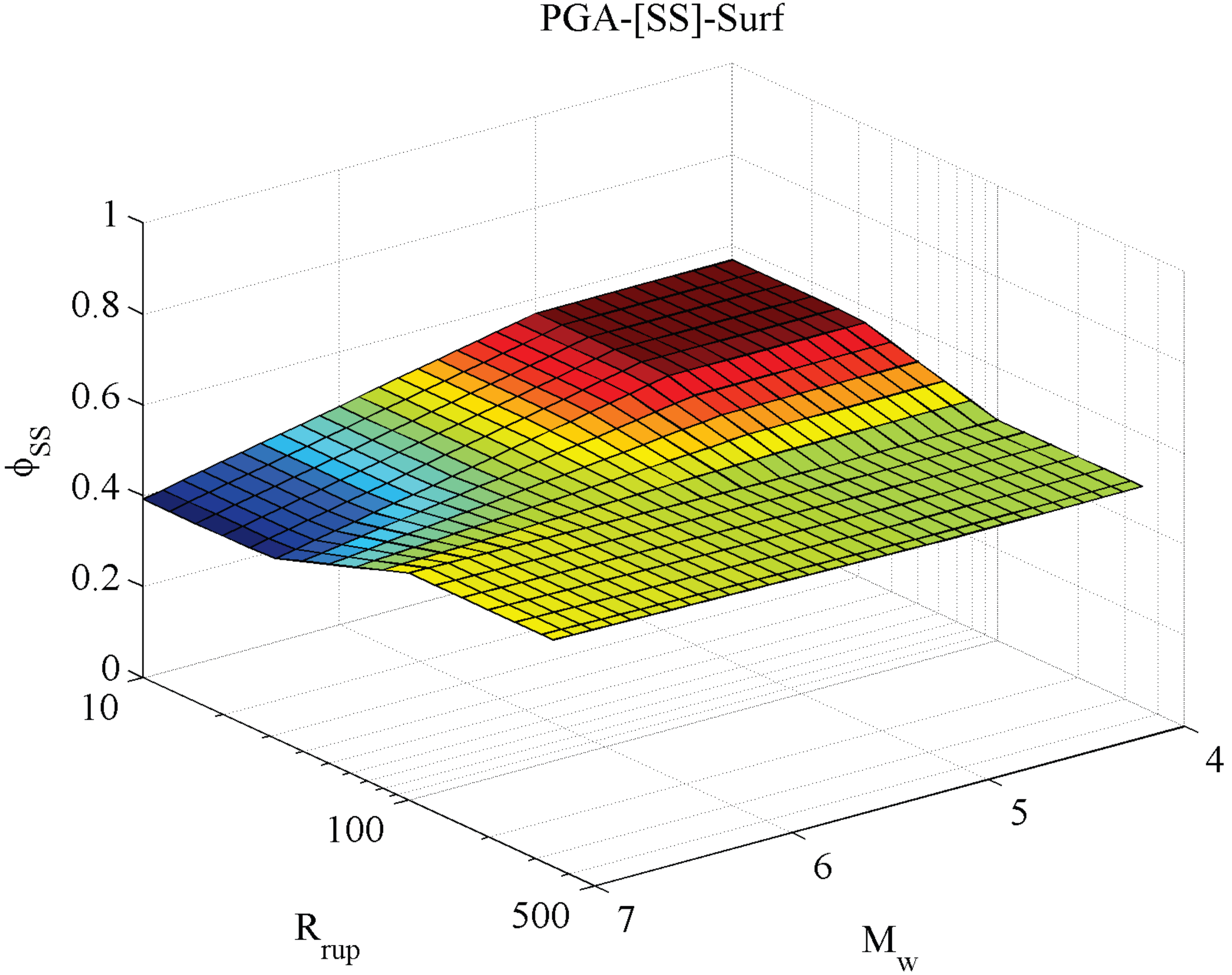
- Issues that must be addressed
  - Computation of aleatory standard deviations [$std(\delta_o) = \phi_{ss}$ and $std(\delta B_e) = \tau$]
  - Estimates of site term (δS_s) and its uncertainty

Issues in the computation of aleatory variability

Estimate of σ_{ss}

$$\sigma_{ss} = \sqrt{\tau^2 + \phi_{ss}^2}$$

- ϕ_{ss} has been shown to be relatively stable regionally and across tectonic regimes
    - ϕ_{ss} has also been shown to be magnitude dependent: differences in ϕ_{ss} between studies tend to be associated to the range of magnitudes considered in the data



From Dawood (2014)

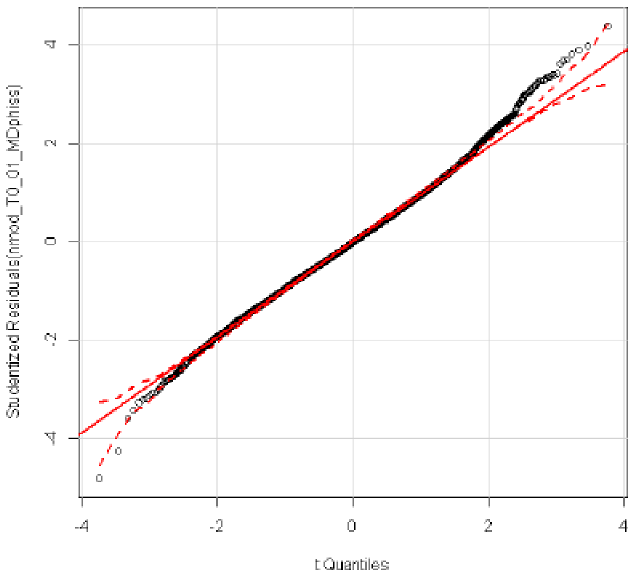
Estimate of σ_{ss}

$$\sigma_{ss} = \sqrt{\tau^2 + \phi_{ss}^2}$$

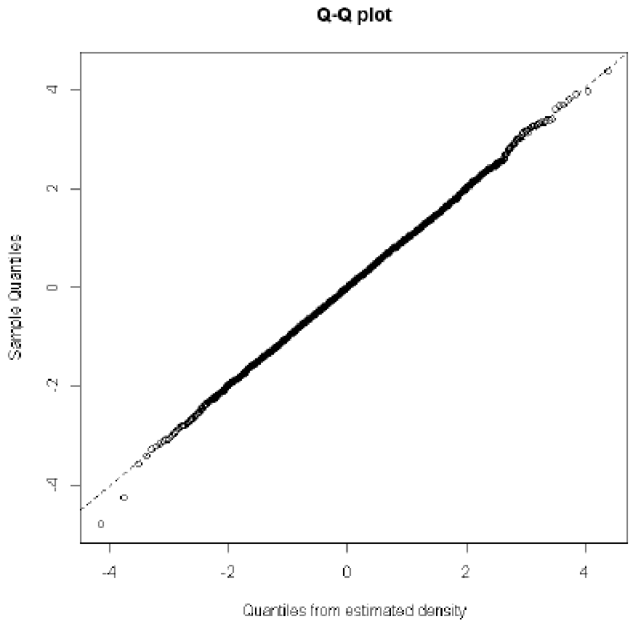
- ϕ_{ss} has been shown to be relatively stable regionally and across tectonic regimes
  - ϕ_{ss} has also been shown to be magnitude dependent: differences in ϕ_{ss} between studies tend to be associated to the range of magnitudes considered in the data
- Residuals have "heavy tails"- use a mixed model

PGA ϕ_{ss} Mag Dep – ASK2014

Q-Q Plot for log normal residuals



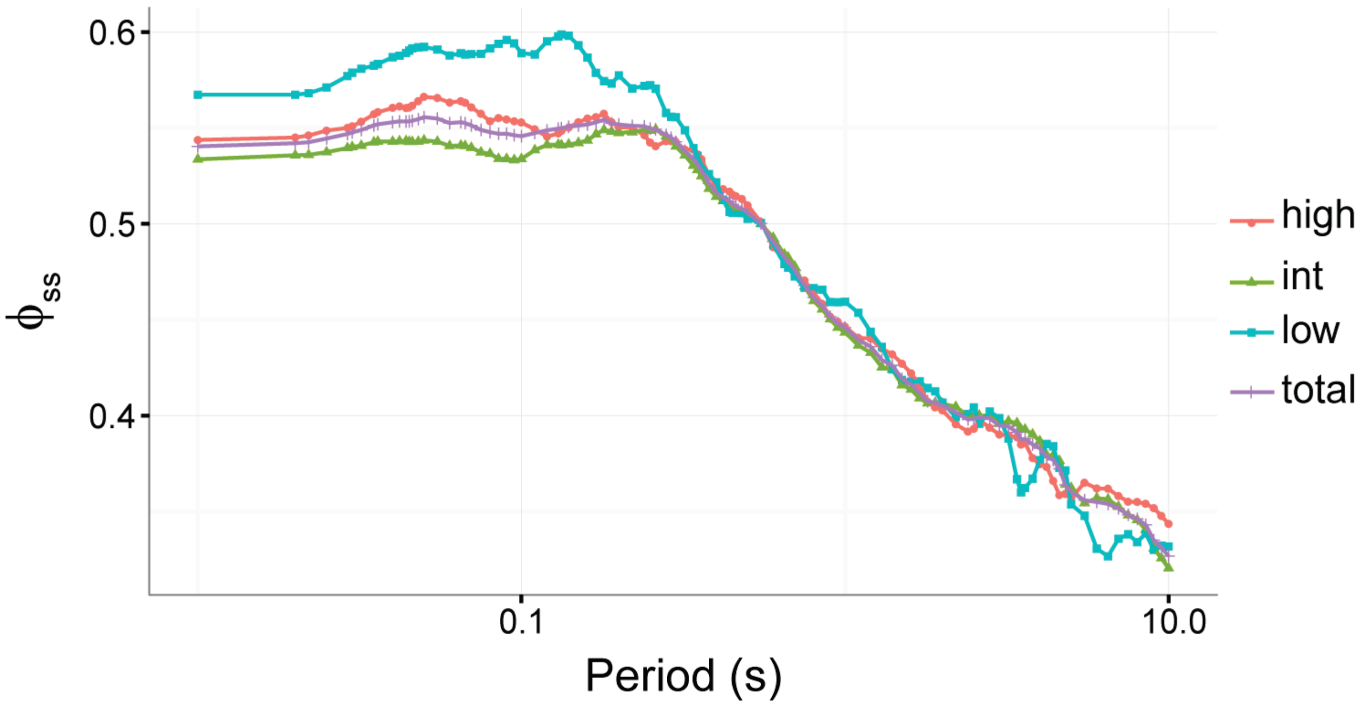
Q-Q Plot for mixture of two log normal s
~ sigma ratios of 1.2 and 0.8, ~ 50%-50%



From HPSHA final report

Estimate of σ_{ss}

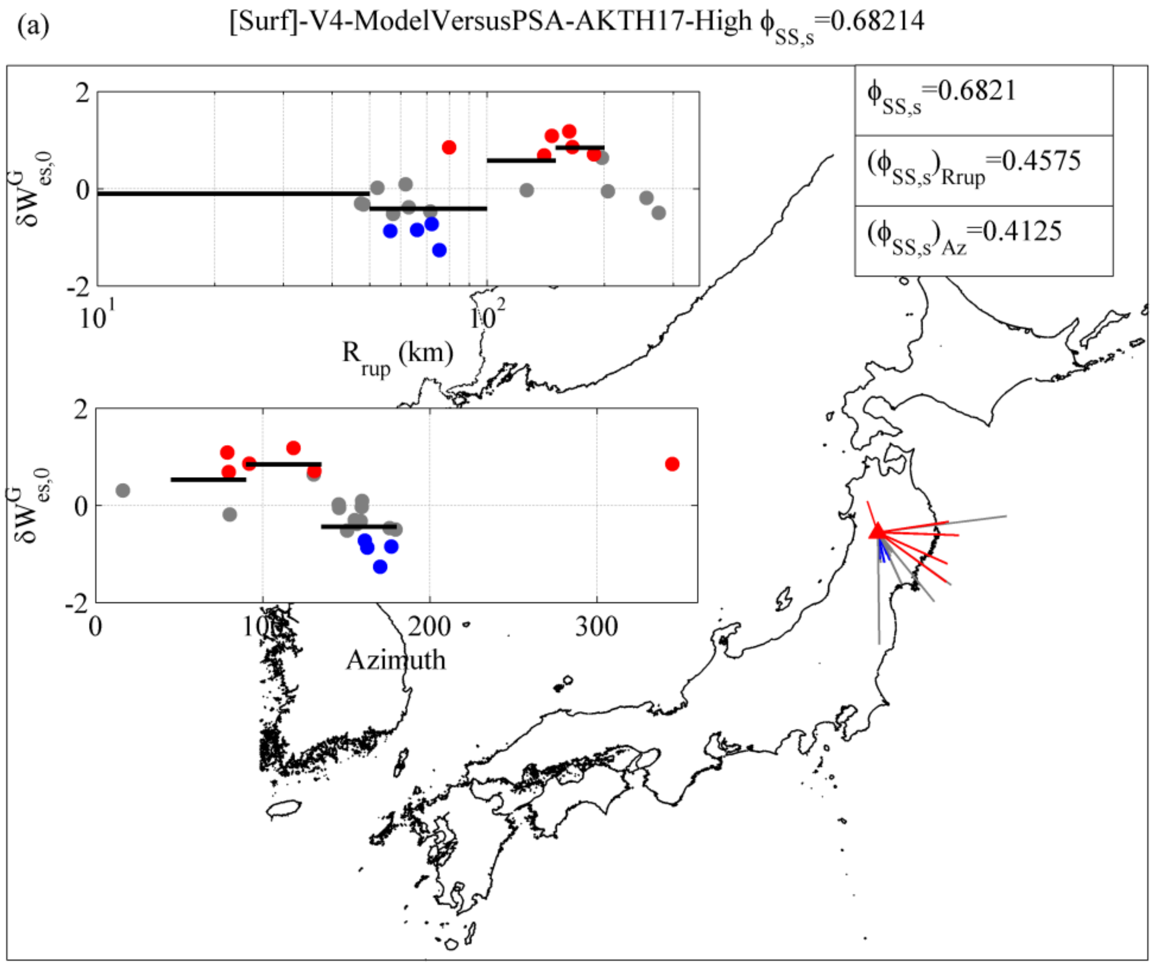
- Certain conditions can lead to higher-than-average values of single-station sigma

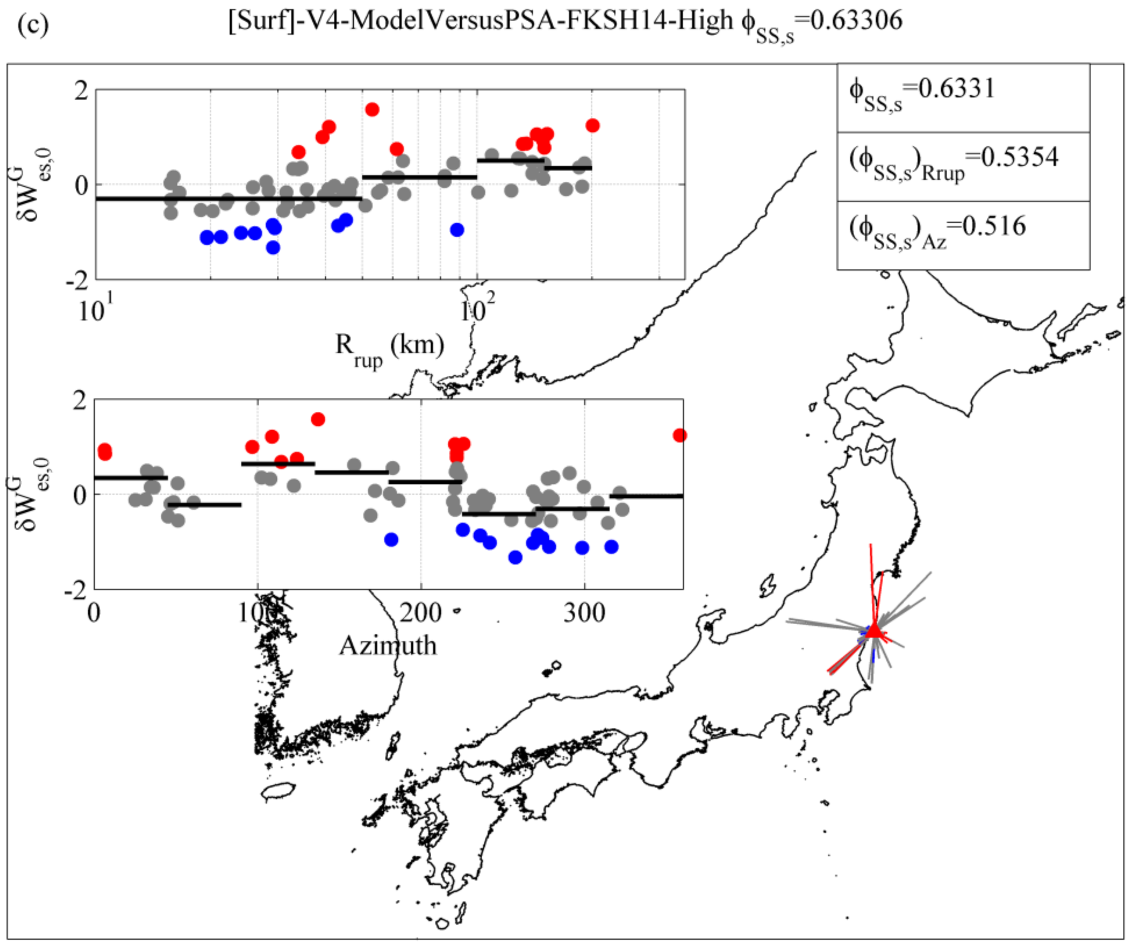


Rai et al. (2015). Differences in single-station sigma estimates due to surface topography

Estimate of σ_{ss}

- Azimuthal stability of site term affects estimates of ϕ_{ss}

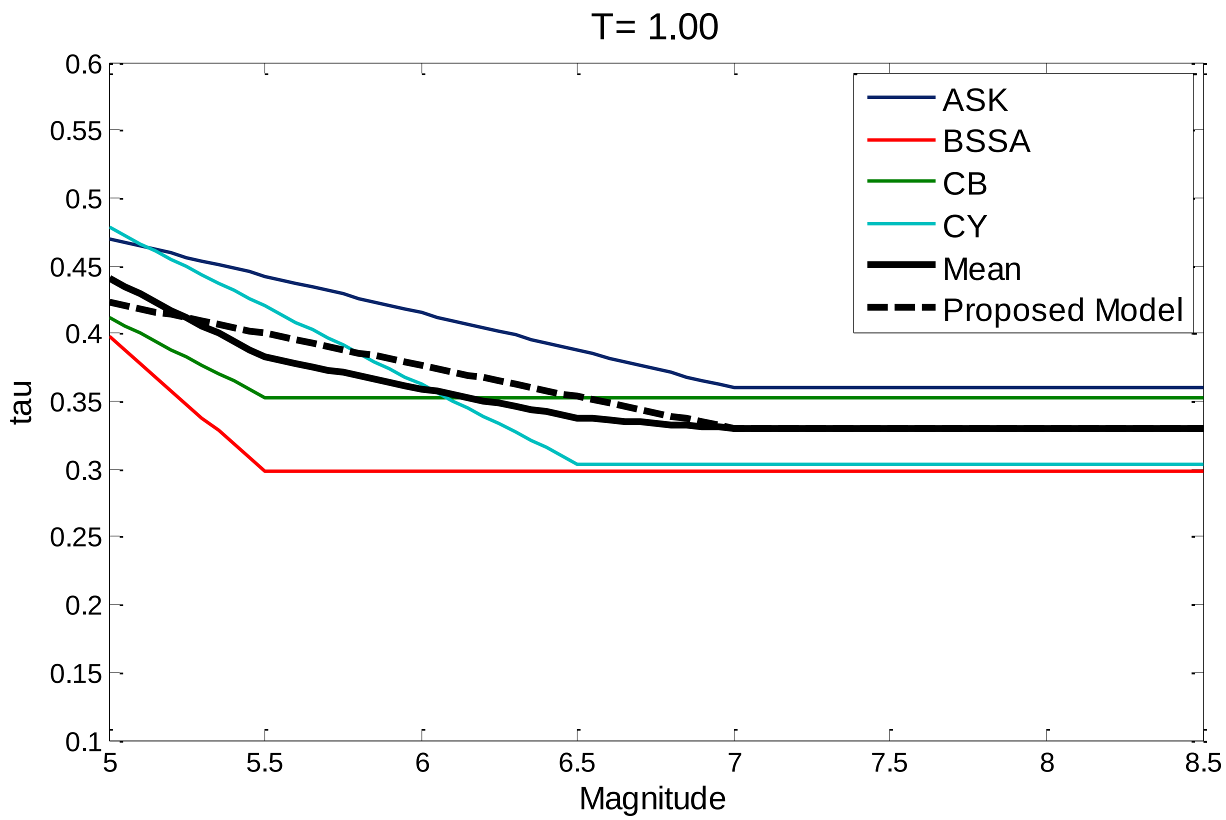




Estimate of σ_{ss}

$$\sigma_{ss} = \sqrt{\tau^2 + \phi_{ss}^2}$$

- Computation of τ
  - Important variations amongst GMPEs – indication of possibly important regional variations
  - Magnitude dependence in most up-to-date models



From HPSHA final report

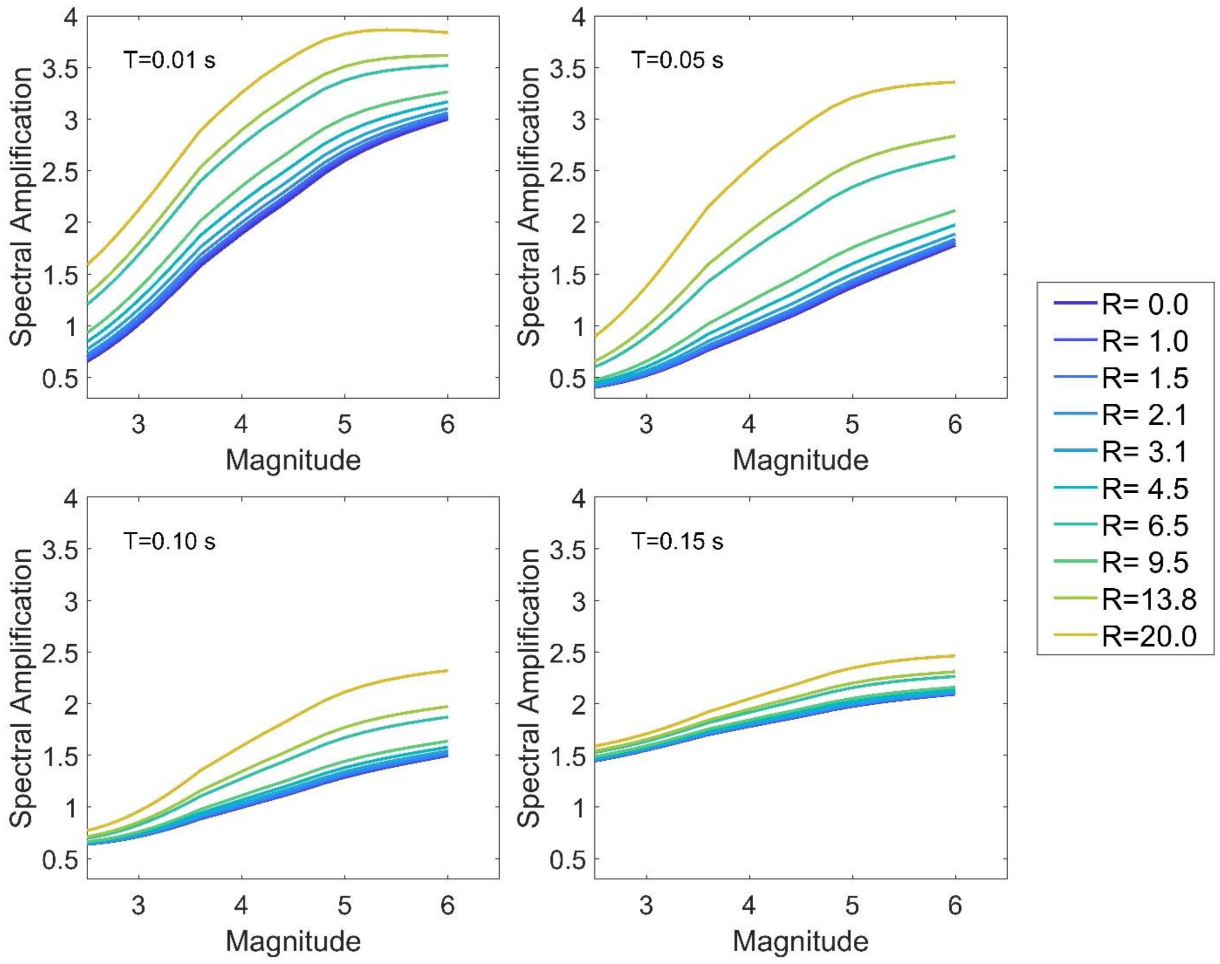
Estimation of the event term and its uncertainty

How to estimate the event term

- In GMPEs: results from the regression
- For sites to be analyzed
  - Empirically: through recordings
  - Analytically: site response exercise

Site term in GMPEs

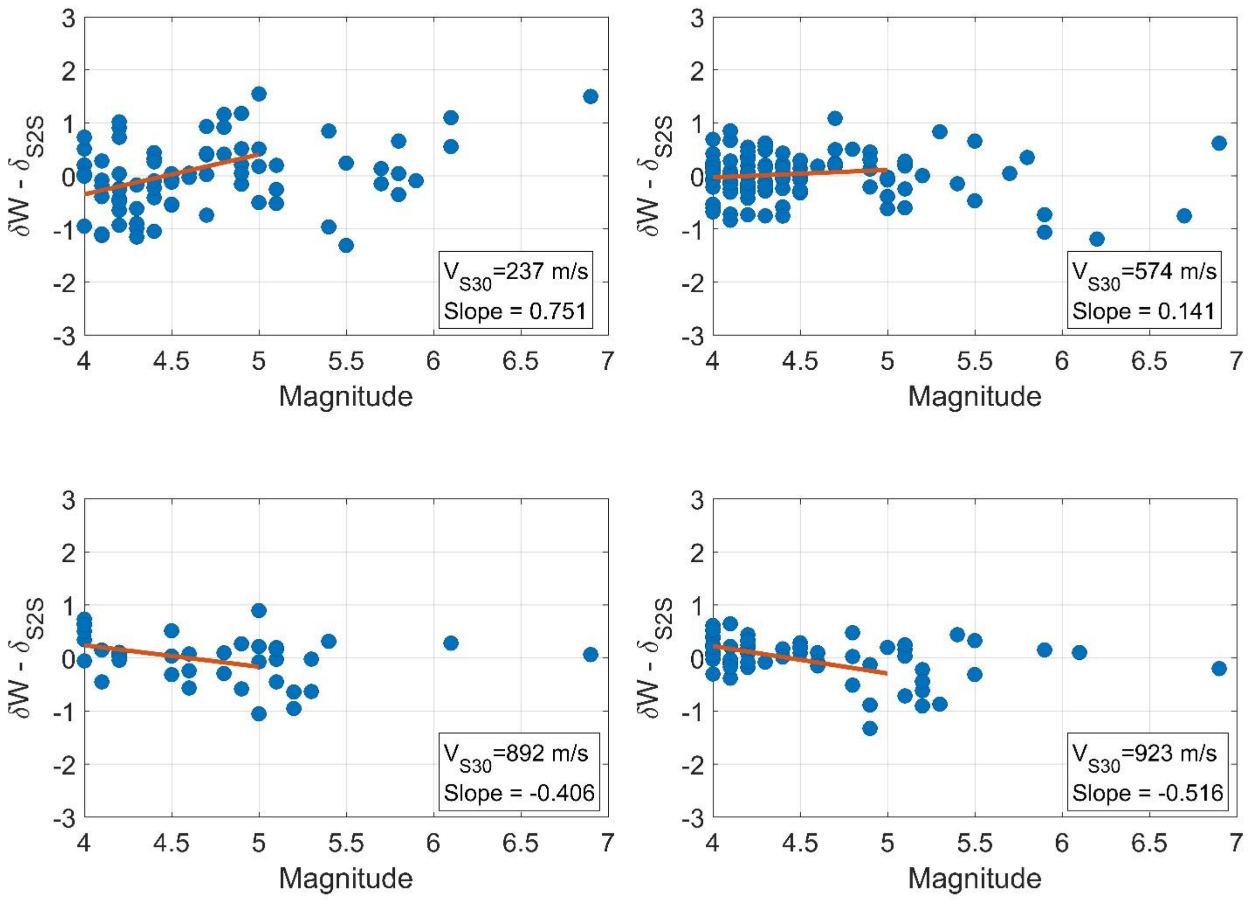
- In most analyses, the site term is assumed to be a constant (i.e., magnitude and distance independent
- However … for sites with strong attenuation, analytical estimates of linear site response show both magnitude and distance dependence at high frequency (in response spectra)



From Stafford et al. (2017)

Site term in GMPEs

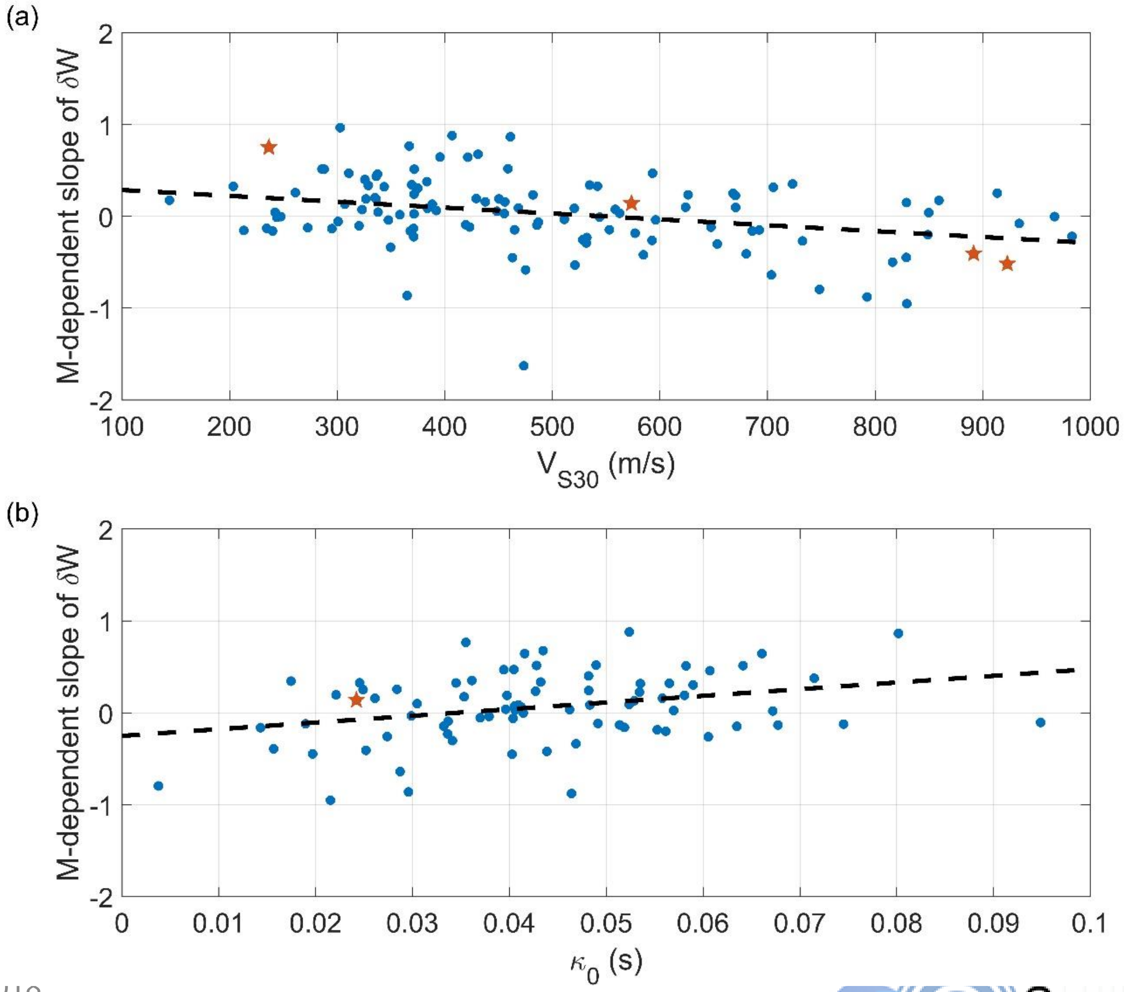
- Magnitude dependence is also seen in site terms from GMPEs



Magnitude dependence also seen in KiK-net event-corrected residuals (from Stafford et al., 2017)

Site term in GMPEs

- Magnitude dependence is also seen in site terms from GMPEs



Magnitude dependence also seen in KiK-net event-corrected residuals (from Stafford et al., 2017)

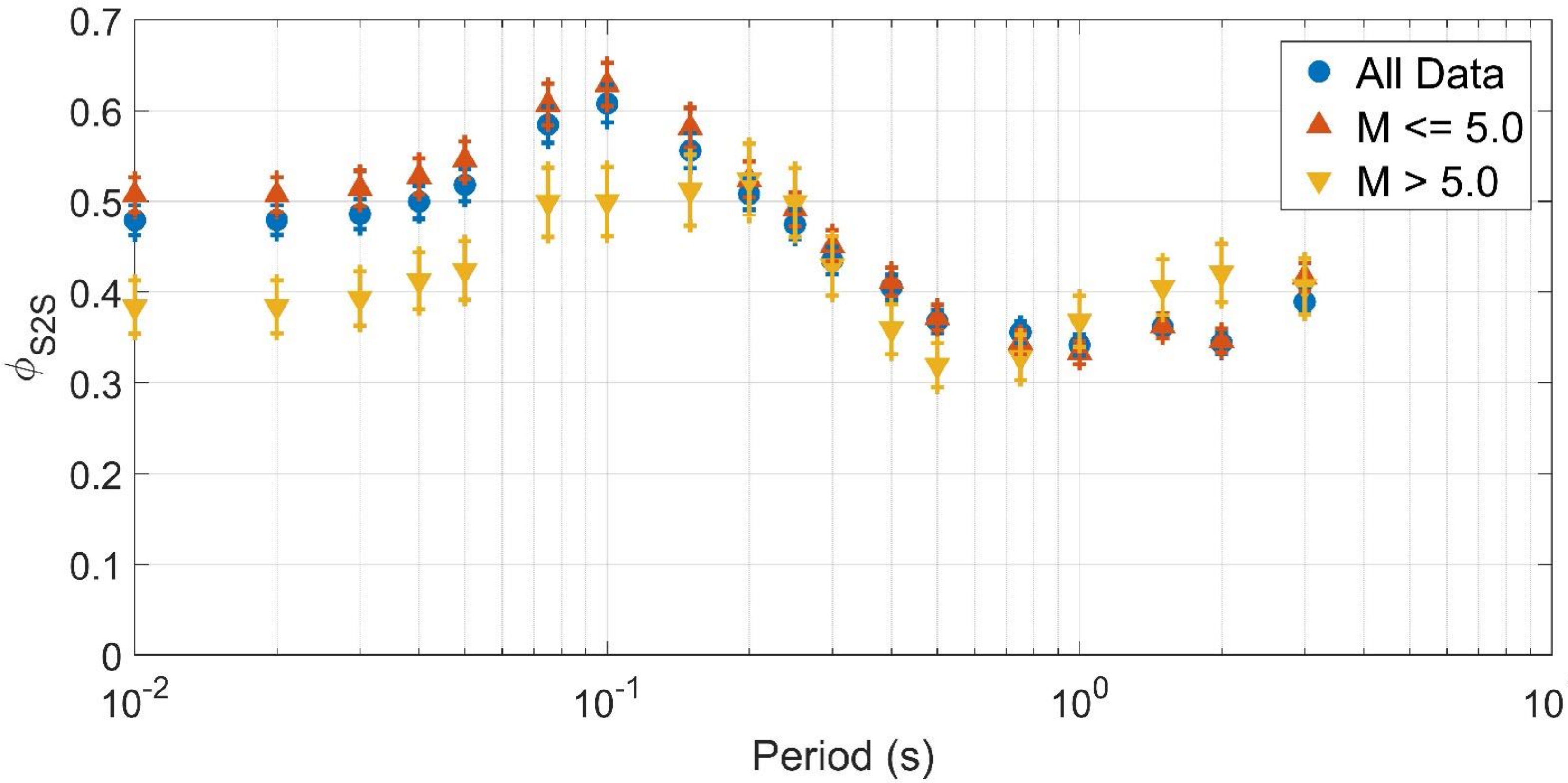
Site term in GMPEs

Implications

- By assuming that the site term is constant, the magnitude-dependence is transferred to the source scaling, leading to erroneous source scaling for small magnitudes
- Event terms estimated with few, small M events can lead to erroneous estimates
    - “updating” of event terms must also consider potential scenario dependence
- Scenario dependence not captured in site terms can pollute M- and R-dependence of other standard deviation components
- Inflating ϕ_{S2S}

Site term in GMPEs

- Scenario dependence leads to magnitude dependence of the site-to-site variability $[std(\delta_S)]$



From Stafford et al. (2017)

Empirical estimates of site term for PSHA

- Obvious challenge: lack of recordings
  - Large statistical uncertainty
  - For short periods: poor extrapolating power beyond magnitude-distance range of recordings
  - Few recordings: path term and site term can't always be separated

Analytical estimates of the site term

Assumptions regarding dimensionality of problem

- 1D, vertical propagation
- 2D, vertical propagation
- 2D, variable incidence angles
- 3D, vertical propagation
- 3D, variable incidence angles

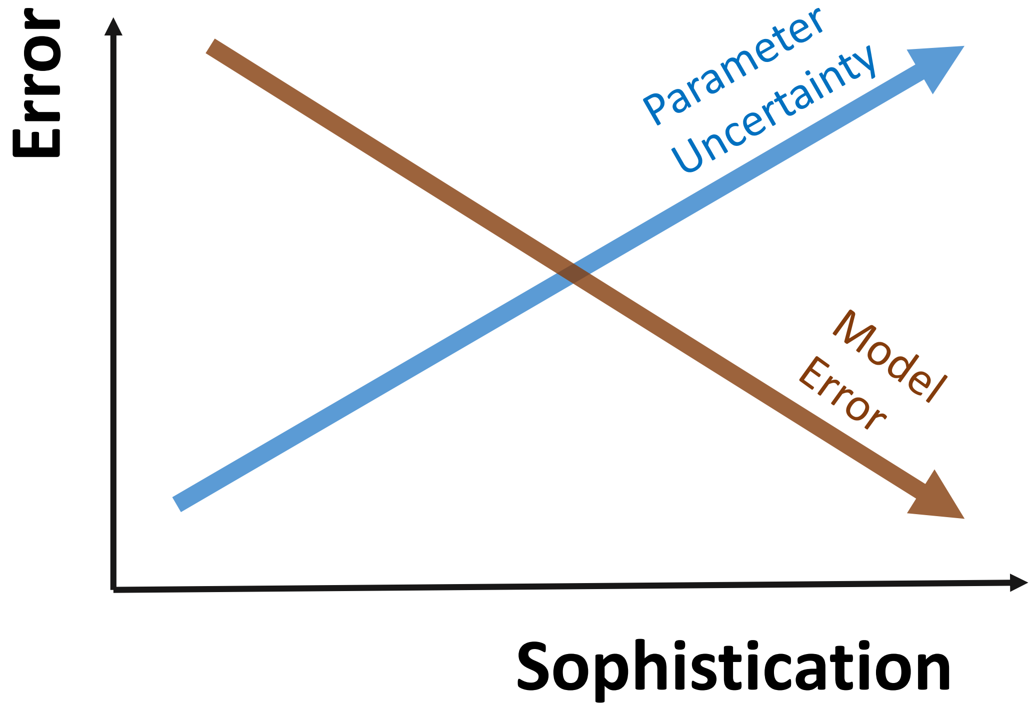
Increased Sophistication

Assumptions regarding material behavior

- Linear
- Equivalent Linear
- Nonlinear, total stress
- Nonlinear, effective stress analysis

Analytical estimates of the site term

- As sophistication increase, the parameterization of the problem increases
- Two relevant types of epistemic uncertainty
  - Model error
  - Parameter uncertainty
- Unless supported by a wider characterization effort
  - Increased parameter uncertainty



Analytical estimates of the site term

- Important to consider a variety of approaches
  - More complex approaches: identify repeatable trends
  - Simpler approaches: appropriately map parameter uncertainty
- It is important to properly quantify error in prediction (epistemic uncertainty of site term)

Groningen site response study: bounds on epistemic uncertainty of site term

- Lower bound: model uncertainty
  - Establish a lower bound to account for model error
  - Based on other studies of downhole arrays
  - Take into account detail of characterization of Groningen field

- Upper bound
  - No more uncertainty than Vs30-based NGA models
  - These models use a simple parameterization using data from multiple regions

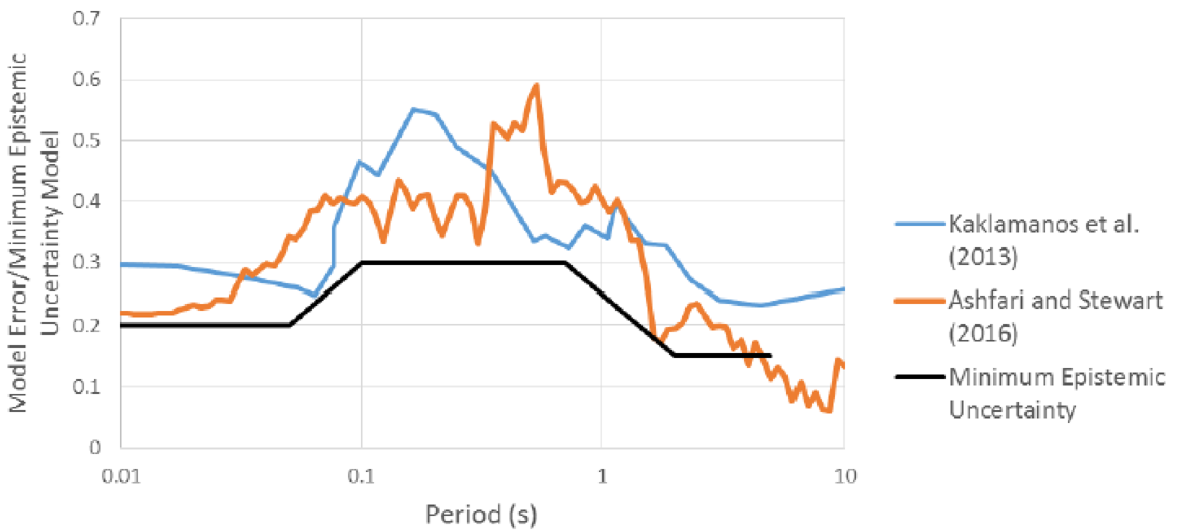


Figure 9.11. Model error computed by Afshari & Stewart (2016c) and Kaklamanos *et al.* (2013), and model for minimum epistemic uncertainty ($\phi_{S2S,min}$)

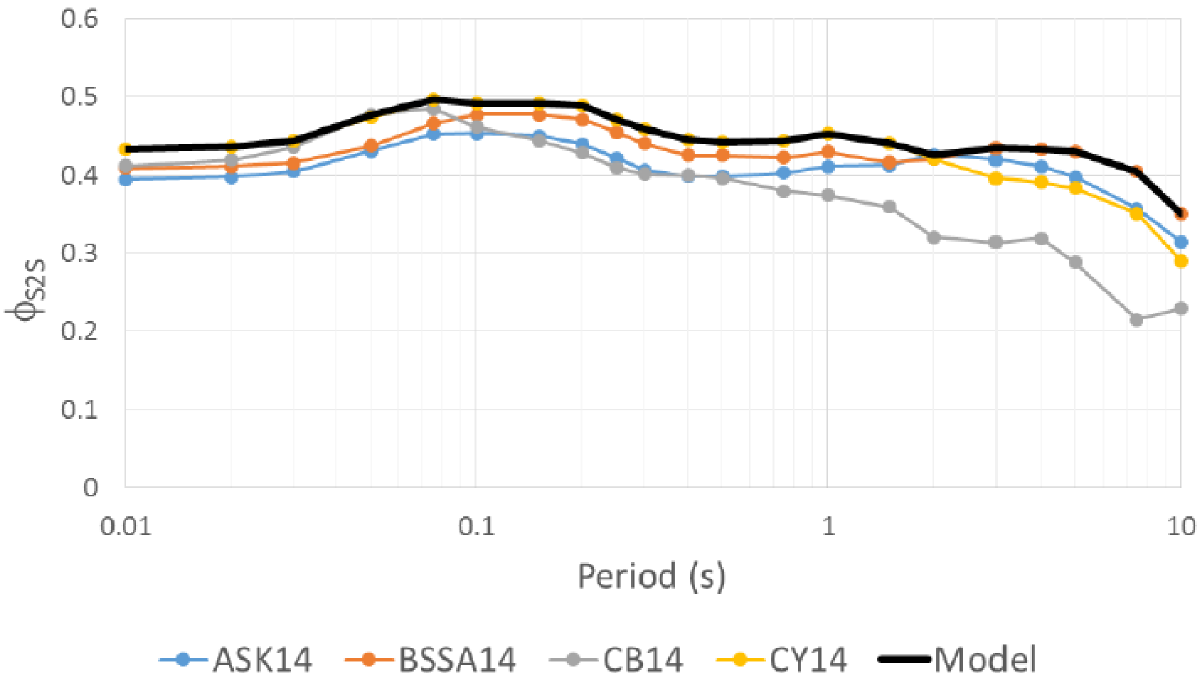
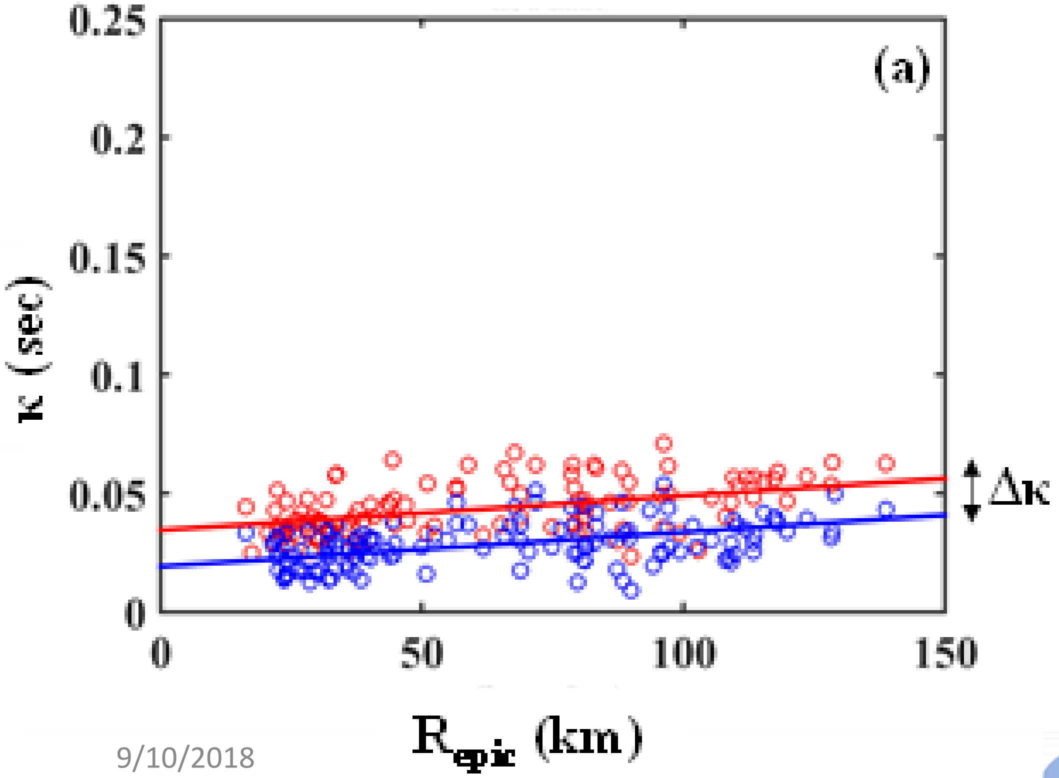


Figure 9.12. Model for the upper bound to ϕ_{S2S}. The values of ϕ_{S2S} for the NGA W2 models were provided by Linda Al Atik (*personal communication*).

Issues to consider for the analytical estimates of the site term:Damping

- Site response predictions require an estimate of material damping
- Generally, geotechnical models based on laboratory data are used
- Studies of vertical arrays indicated that laboratory models underestimate field damping

Station IBRH12 (KiKnet)



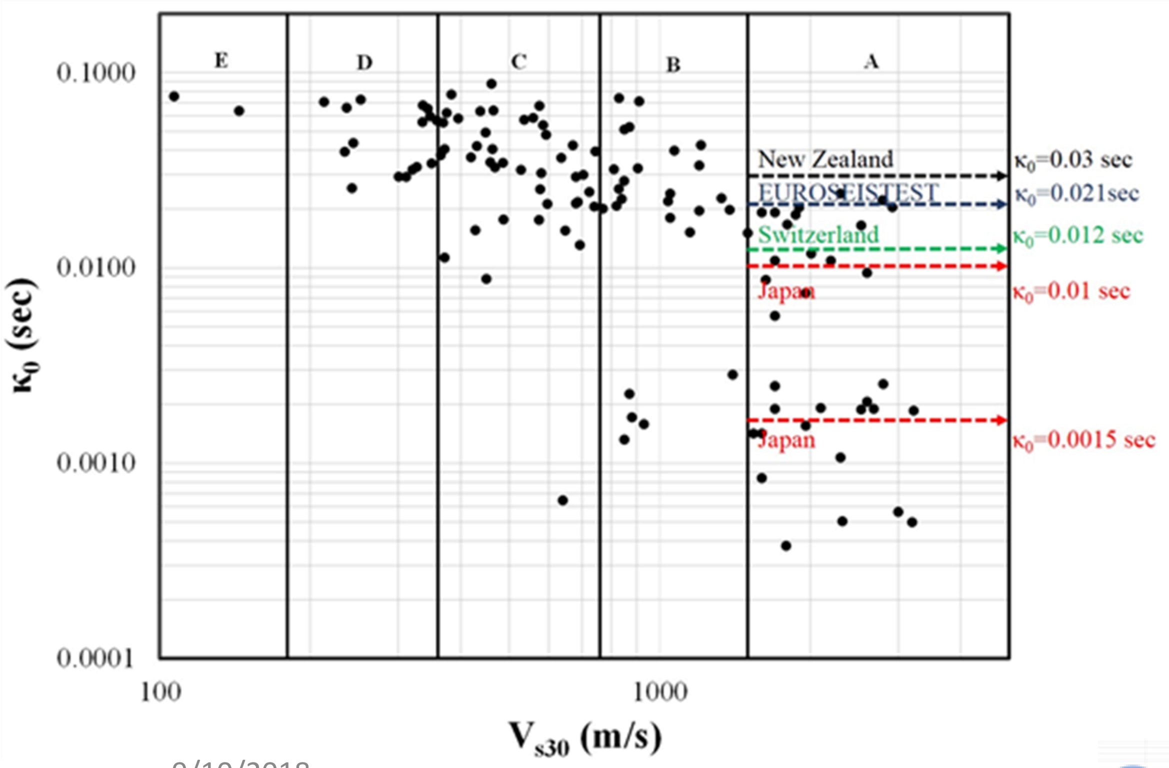
$$(\xi_{min})_{new} = (\xi_{min})_{lab} + \Delta\xi_{min}$$

Using all KiKnet surface and downhole recordings, along with “reasonable” estimates of geotechnical laboratory damping, $\Delta\xi_{min}$ values from 2% to 3% are needed to obtain low strain damping profiles which are compatible with the measured κ_0 at the selected sites.

From Cabas et al. (2017)

Issues to consider for the analytical estimates of the site term: Damping

- Site response predictions require an estimate of material damping
- Generally, geotechnical models based on laboratory data are used
- Studies of vertical arrays indicated that laboratory models underestimate field damping



$$(\xi_{min})_{new} = (\xi_{min})_{lab} + \Delta\xi_{min}$$

Using all KiKnet surface and downhole recordings, along with "reasonable" estimates of geotechnical laboratory damping, $\Delta\xi_{min}$ values from 2% to 3% are needed to obtain low strain damping profiles which are compatible with the measured κ_0 at the selected sites.

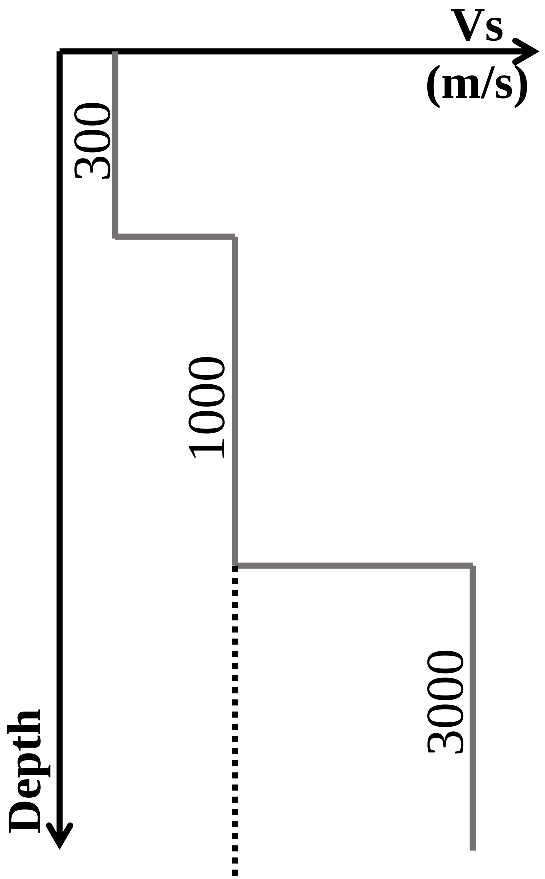
ΣIGMA2

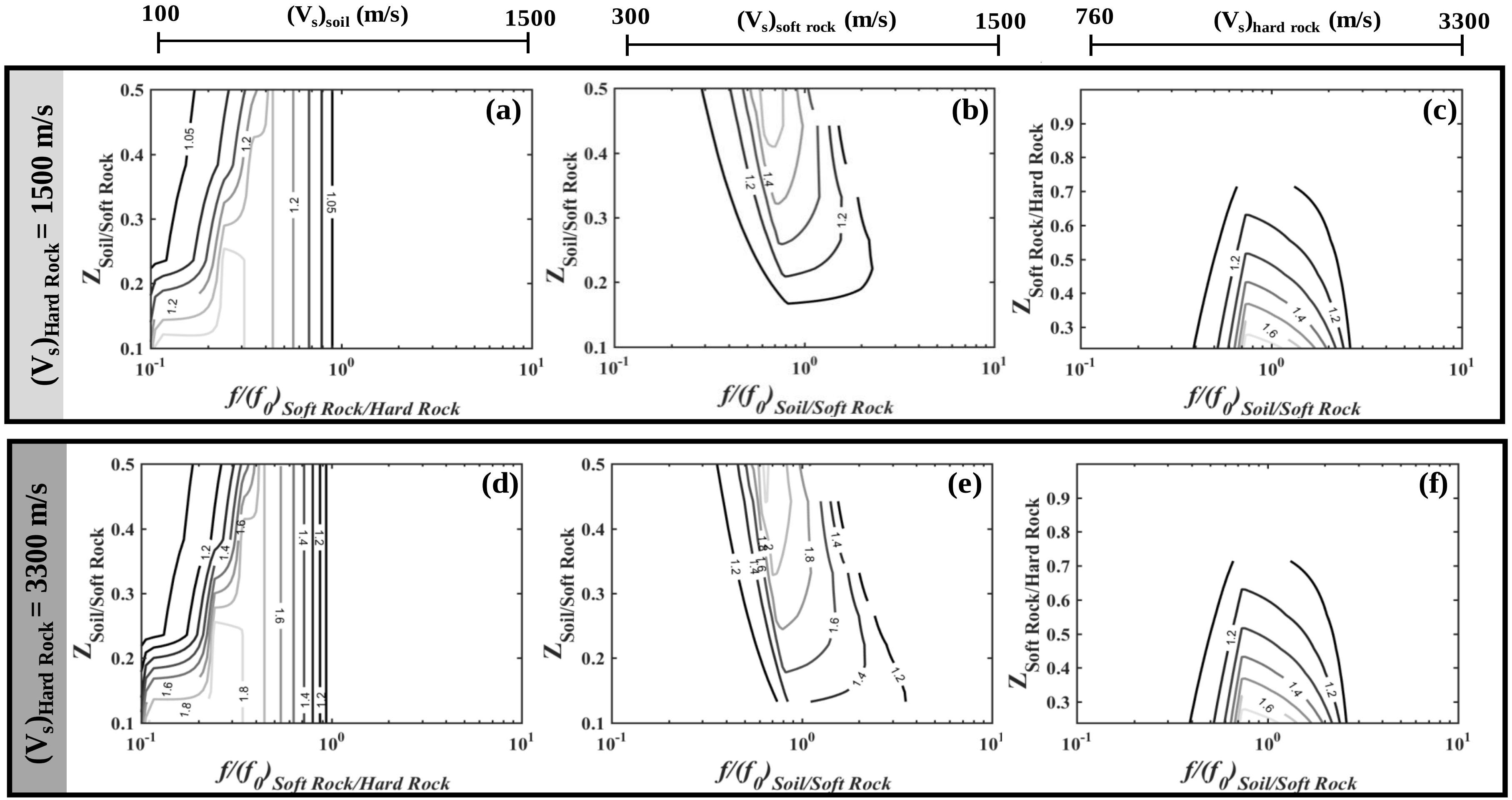
From Cabas et al. (2017)

Issues to consider for the analytical estimates of the site term: Location of Elastic Half-Space (EHS)

VirginiaTech Invent the Future

Error in Transfer Function associated with erroneous selection of the EHS (from Cabas et al., 2018)





Implementation Issues

Implementation Issues

In site response, analytical estimates are expressed in terms of Amplification Factors (AF)

$$\ln(AF) = f_1^* + f_2 \ln\left(\frac{Sa_{NS_B,g}+f_3}{f_3}\right) + \varepsilon\sigma_{lnAF}$$

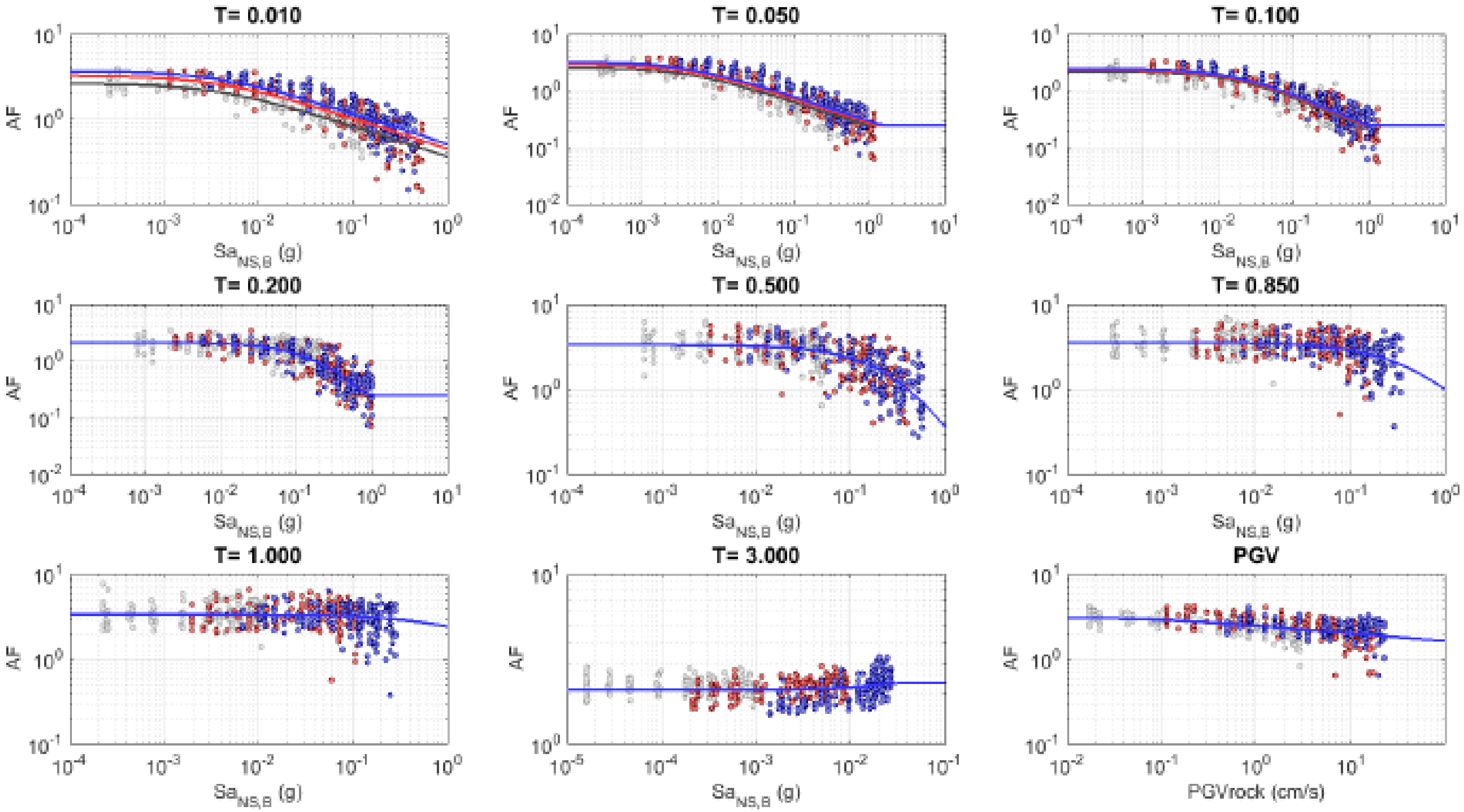
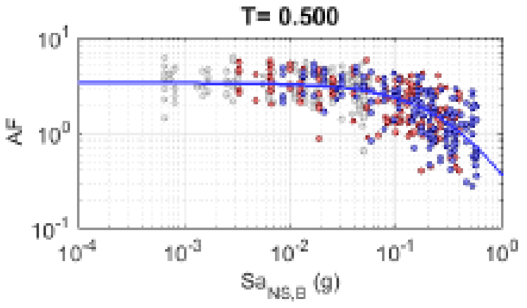


Figure 8.30. Amplification factors (AF) for zone 1208 for selected periods. The colours represent M=3.8 (black), 4.8 (red) and 5.8 (blue). The fit through the data is represented by the solid line, the fit line is computed for a single distance (R=5 km). S_{aNS_B} has units of g.

Implementation Issues

$$lnAF = \overline{lnAF} + \sigma_{lnAF}$$

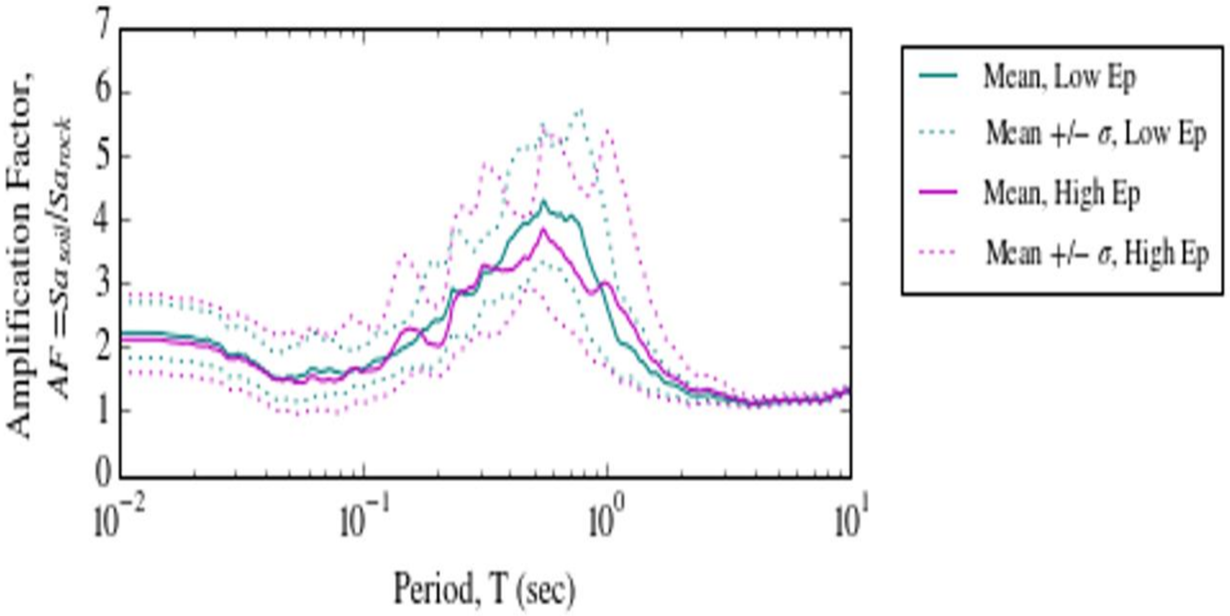


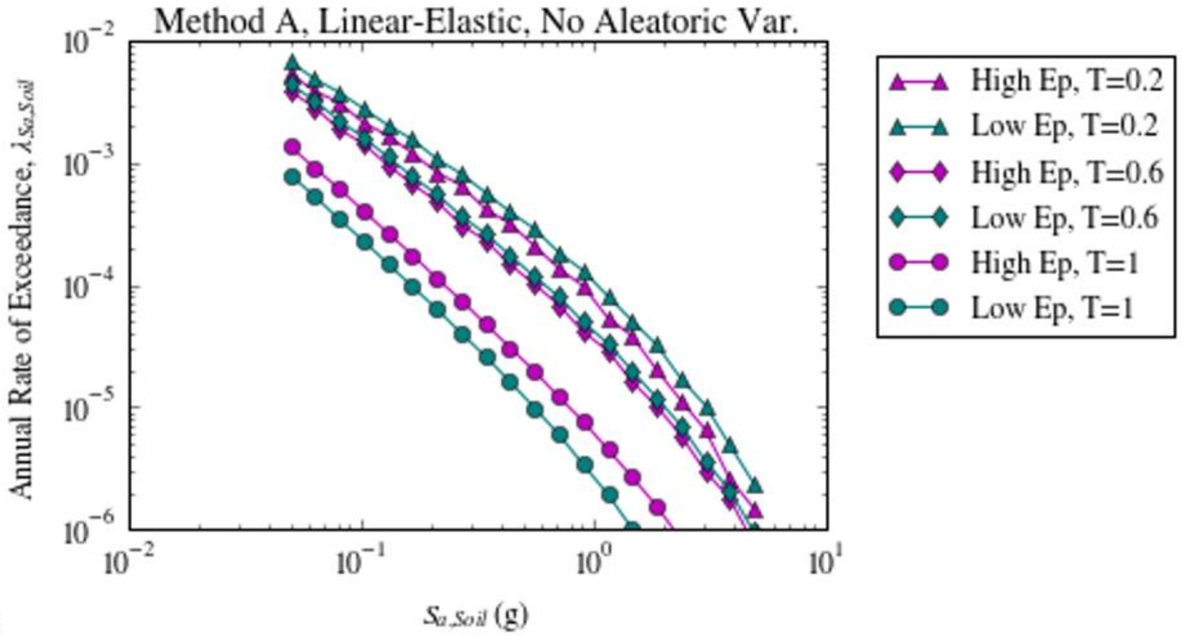
- σ_{lnAF} includes (depending on the implementation)
  - Motion to motion variability
  - Aleatory variability in Vs profile (assuming that proper partition can be made)
- Question: should this uncertainty be included in hazard computation? Should it be added to the epistemic uncertainty of the site term?

- Model error in site amplification predictions maybe correlated to uncertainties in input parameters to site response

Implementation Issues

- There is the possibility that high-epistemic uncertainty smooths out the peaks of the AF curve
  - The impact is reduced when integrating across uncertainty





Courtesy of R. Green and K. Ulmer

Proposed Sigma2 Research at Virginia Tech

- Task 1: Develop GMPEs with proper site-term scaling
  - Expand database of Dawood et al. (2016)
  - Develop a GMPE with proper site scaling
- Task 2: Study the stability of the site term
  - Across tectonic regions
  - Across azimuthal distributions
  - Include effects of nonlinearity?
- Task 3: Establish robust estimates of the model error of 1D site response analysis
  - Use KiKnet surface/borehole recordings (expand work of Thomas and Kaklamanos)
  - Explore other downhole arrays (Stewart et al. 2016)

Proposed Sigma2 Research at Virginia Tech

- Task 4: Determine proxies that can identify a-priori candidates for 1D site response
  - H/V ratios
  - Stability of H/V ratios across rotations
  - Geological/topographical parameters
- Task 5: Implementation in PSHA
  - Correlations of uncertainties
  - Impact of epistemic uncertainty

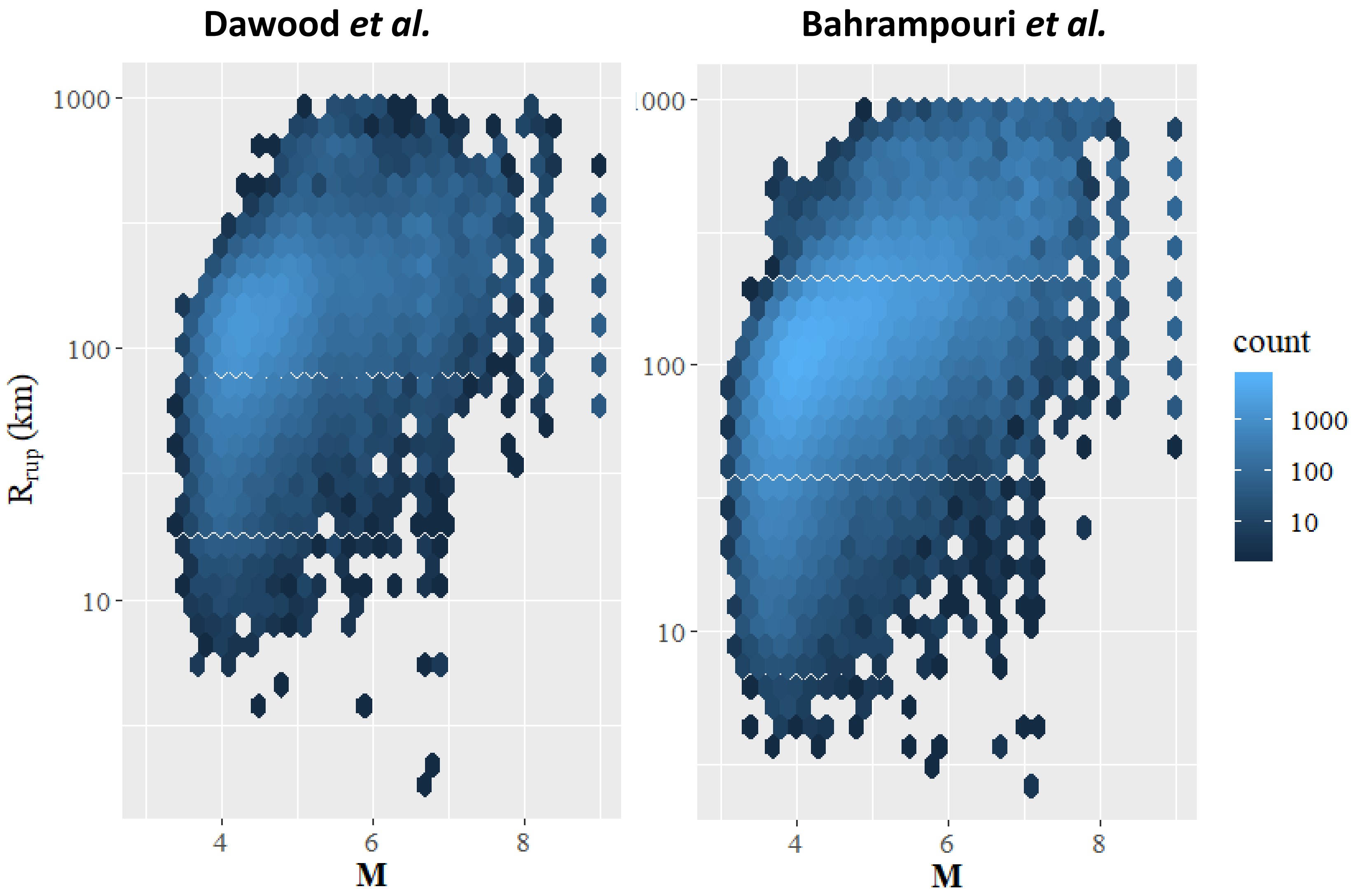
Results to date: Updated Database

Updates from Dawood et al. (2016; data up to 2011) to data up to 2017

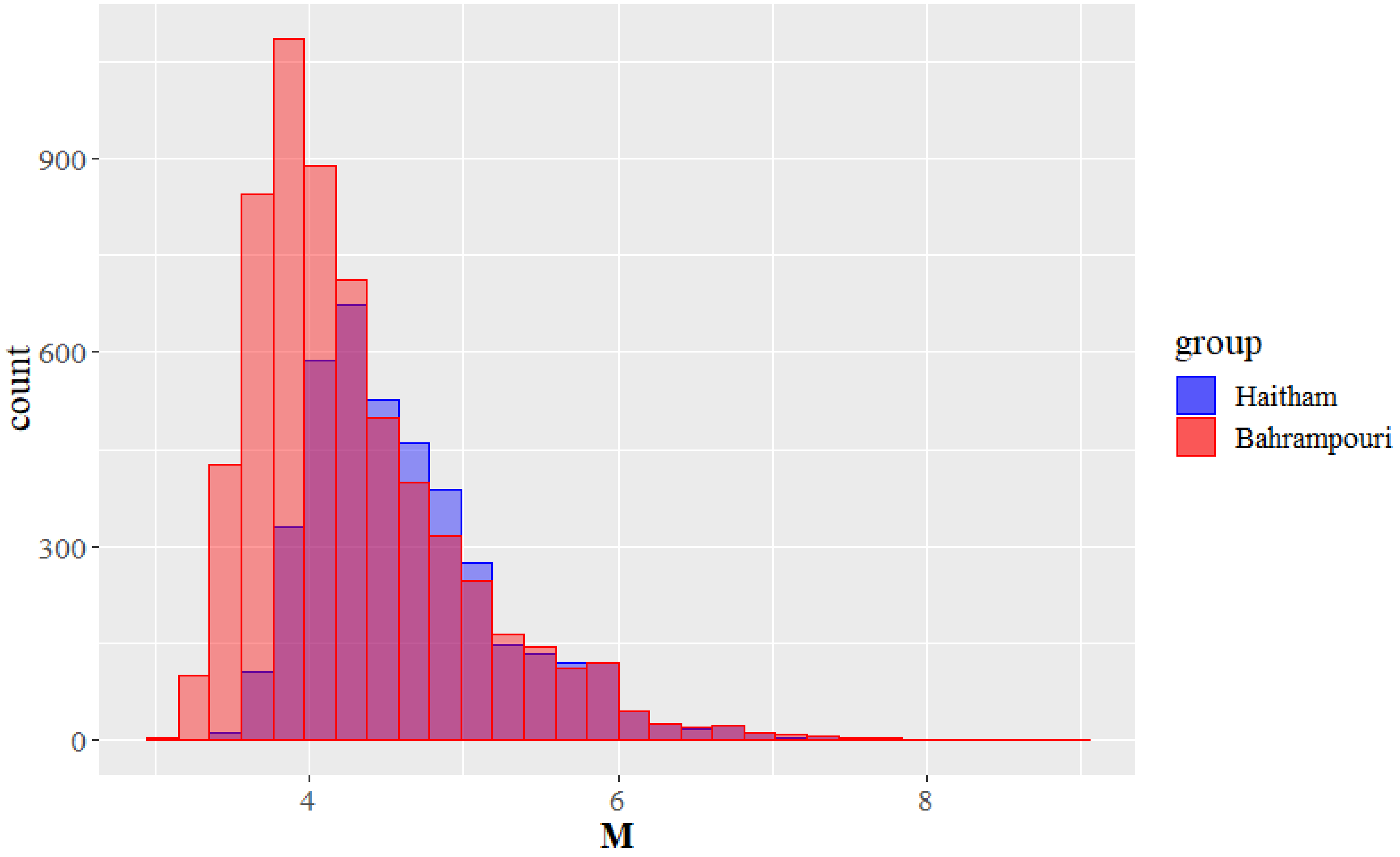
Available data

GM metadata and GM parameters (for each direction)		Event catalog	Site catalog
• Record date • Origin time • Recording duration • Origin date • Sampling frequency • Ratio of usable band pass • Filter low and high and low corner frequency • Spectral accelerations • Fourier amplitude	• PGA • Arias Intensity • D_{S5_75} • $D_{S5\_75\_sorted}$ • D_{S20_80} • R_{epi} • R_{hypo} • R_{jb} • R_{rup} • R_x • R_y • R_{y0} • U (directivity parameter) • T (directivity parameter)	• Magnitude based on Fnet and KiK-net • Depth of earthquake based on Fnet and KiK-net • Tectonic regime based on Garcia and Zhao algorithms • Focal mechanisms based XX and YY • Magnitude based on published models	• Vs30 • Vsx (x=5, 10, 30, 50, 100 m) • Z1 • H800 • Latitude • Longitude • Depth to bottom of borehole • Vs at the bottom of borehole

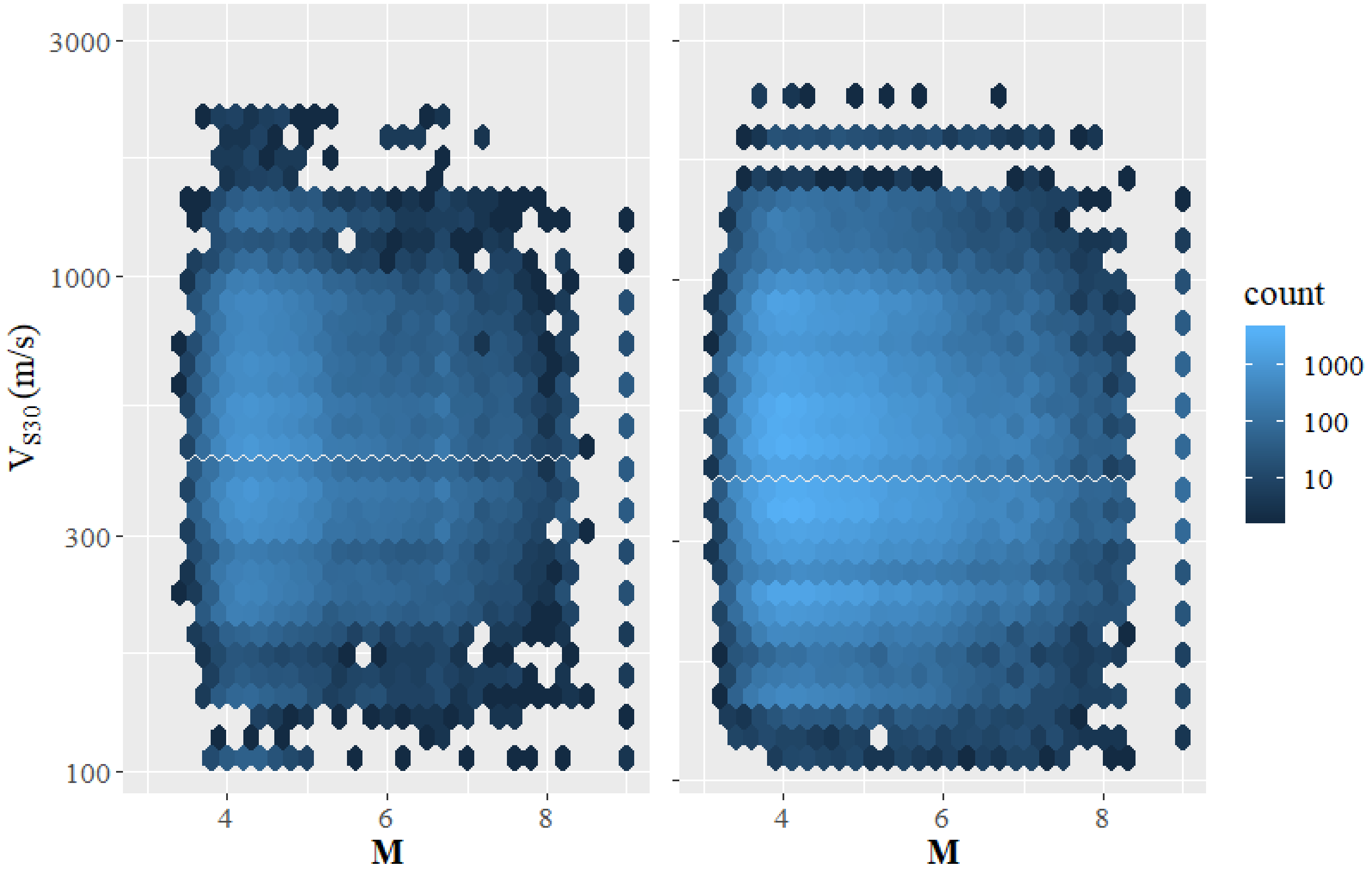
Updates on Database



Plots include all records with moment tensor solutions and acceptable signal to noise ratio


count
900
600
300
0
4
6
8
M
group
Haitham
Bahrampouri

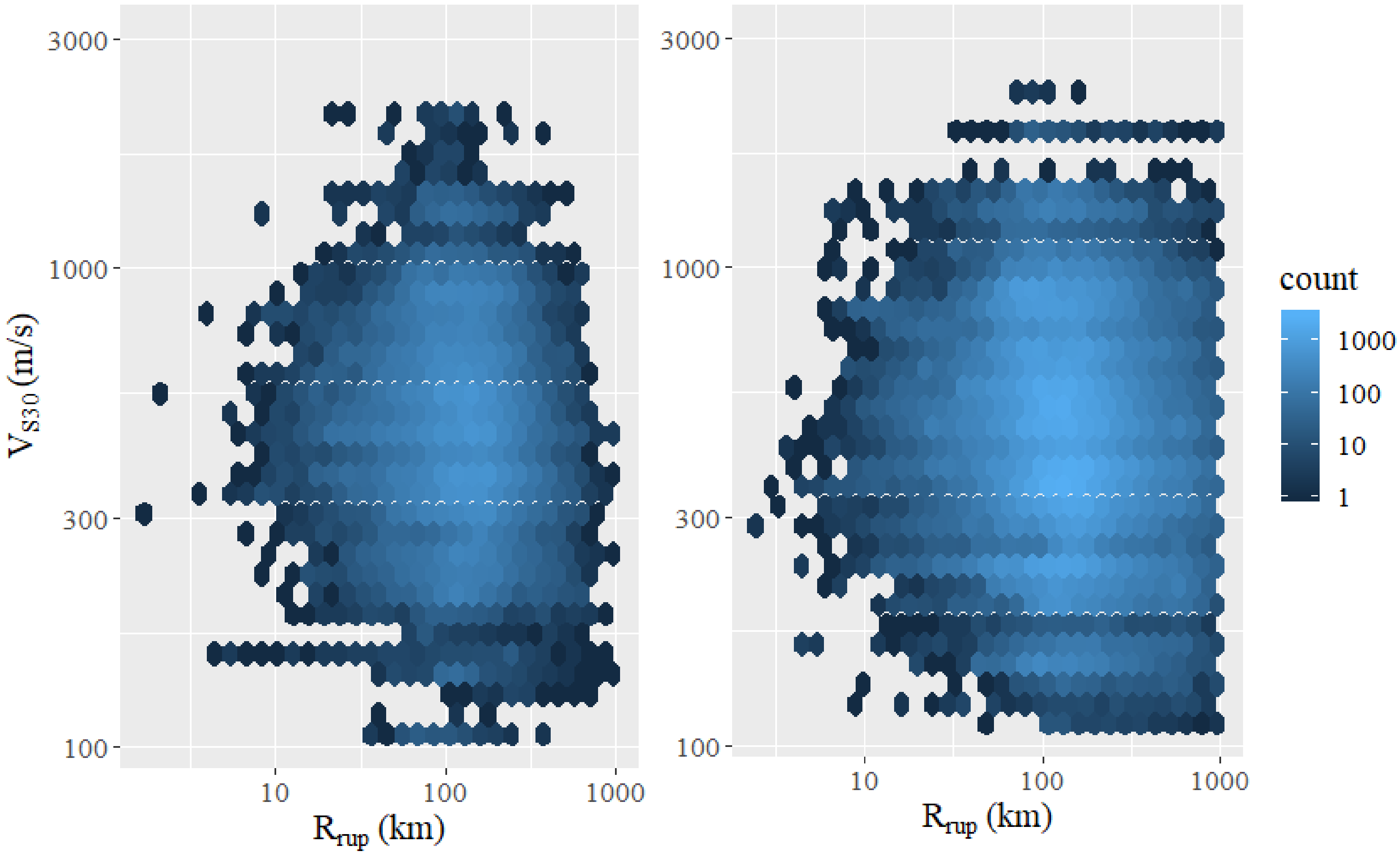
Updates on Database



Dawood *et al.* (2015): Up to 2011
Bahrampouri *et al.* (2018): Up to 2017

Comparison

Dawood *et al.* **Bahrampouri *et al.***



#	Event	Time	M	Number of recording	#	Event	Time	M	Number of recording
1	**Central Tottori**	**2016**	**6.2**	**269**	14	Miyagi-ken Hokubu	2003	6.4	193
2	Fukuoka1	2005	6.64	265	15	Niigata-Ken Chuetsu	2004	6.62	286
3	Fukuoka2	2005	5.57	131	16	Niigata-ken Chuetsu-oki	2007	6.6	305
4	Fukushima Hamadori	2011	6.68	356	17	Northern Coast, Iwate	2008	6.8	376
5	Geiyo	2001	6.68	244	**18**	**Northern Nagano**	**2014**	**6.7**	**260**
6	Hida Swarm EV09	1998	5.13	13	19	Noto Hanto	2007	6.73	334
7	Honshu(OFF-Miyagi)	2005	7.5	391	20	Off Miyagi(2003)	2003	7	382
8	Iwate	1998	6.27	1	21	Offshore Honshu (Sanriku-Oki)	2011	7.3	185
9	Iwate- Miyagi Nairiku	2008	7	325	**22**	**Offshore Kii Peninsula**	**2016**	**5.9**	**351**
10	Kii Peninsula event 1	2004	7.3	337	23	Suruga bay	2009	6.5	310
11	Kii Peninsula event 2	2004	7.5	361	24	Tohoku	2011	9	525
12	**Kumamoto1**	**2016**	**6.14**	**146**	**25**	Tokachi-oki	2003	8.2	299
13	**Kumamoto2**	**2016**	**7.2**	**328**					

ΣIGMA2

Storage format

- The storage format is hdf5
- HDF stands for Hierarchical Data Format
- HDF format is designed to store and organize large amounts of data
- The data is stored by python
- It has interfaces in many languages including:
  - C, C++, Fortran, Fortran 90
  - Java
  - MATLAB, Scilab or Octave
  - Python supports HDF5 via h5py and via pytables
  - R offers support in the rhdf5 and h5 packages.

VirginiaTech
Invent the Future®

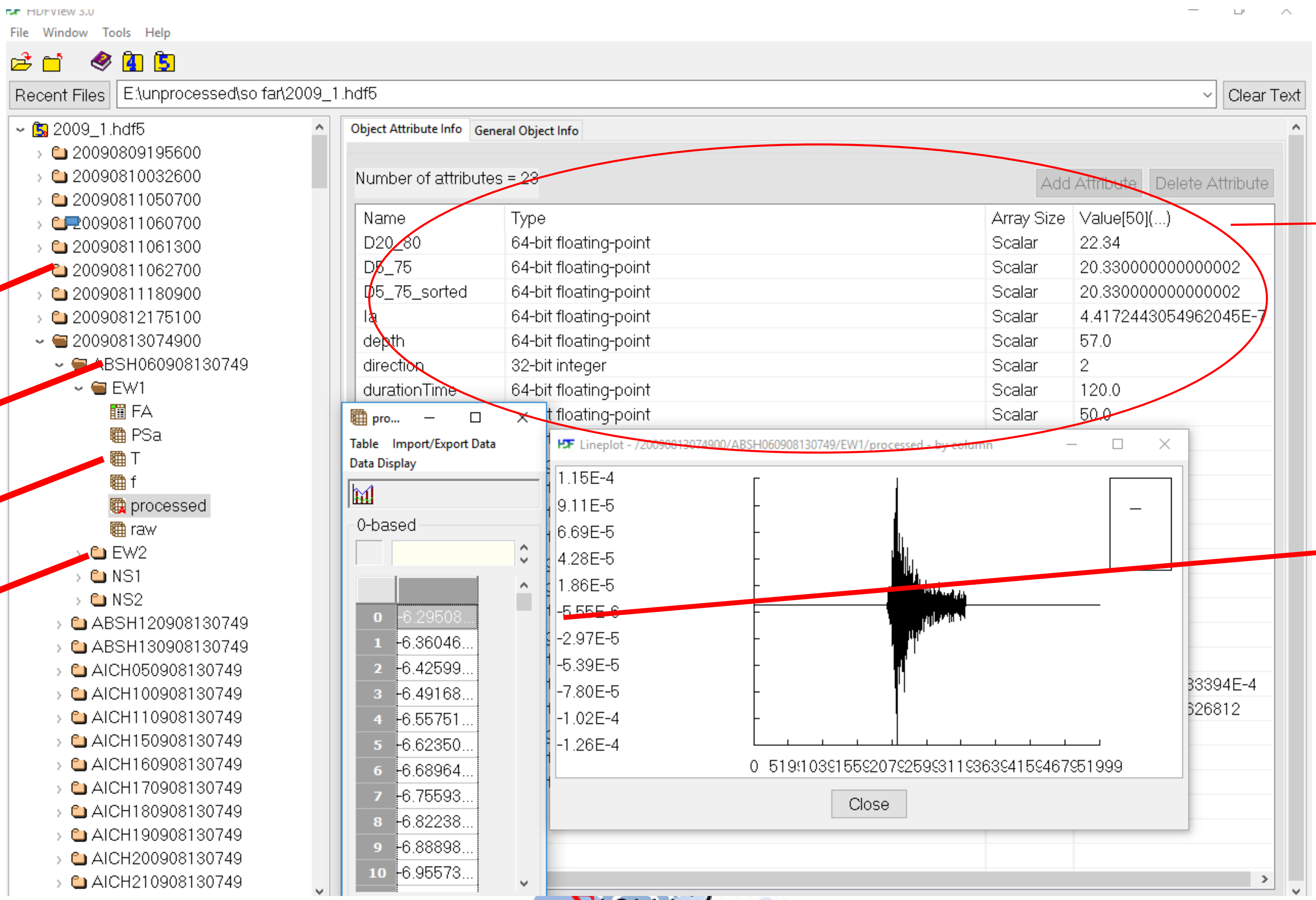
How it look likes

Earthquake

Record

Data

Direction



Attribute

Data

ΣIGMAZ

Thank You